



TITLE:

毎年・非毎年時系列資料を用いた
水文極値の頻度解析に関する研究(
Dissertation_全文)

AUTHOR(S):

田中, 茂信

CITATION:

田中, 茂信. 毎年・非毎年時系列資料を用いた水文極値の頻度解析に関する研究. 京都大学, 2010, 博士(工学)

ISSUE DATE:

2010-01-25

URL:

<https://doi.org/10.14989/doctor.r12423>

RIGHT:

毎年・非毎年時系列資料を用いた水文極値の 頻度解析に関する研究

田 中 茂 信

目 次

1. 緒論	1
参考文献	4
2. 水文頻度解析と水工計画・河川整備	5
2.1 わが国の河川整備と水文頻度解析の歴史	5
2.2 わが国の水文資料の現状と課題	9
2.3 結論	11
参考文献	13
3. 水文頻度解析の手法	15
3.1 頻度解析の一般的手順	15
3.2 水文頻度解析によく用いられる確率分布モデル	16
3.3 推奨される母数推定法	30
3.4 候補モデルのscreening方法	30
3.5 確率水文量の変動性の評価方法	32
3.6 Mathematicaを用いた水文統計解析プログラム	34
3.7 結論	35
参考文献	36
4. 年最大値系列を用いた水文頻度解析	37
4.1 豪雨極値データの解析	37
4.2 洪水極値データの解析	45
4.3 結論	51
参考文献	52
5. 非毎年値資料を用いた水文頻度解析	53
5.1 非毎年値資料を用いることの意義	53
5.2 豪雨のPOT資料の解析	56
5.3 洪水ピーク流量のPOT資料の解析	60
5.4 POT資料における閾値の選定方法	65
5.5 結論	73
参考文献	74

6. AMS資料とPOT資料の比較	75
6.1 AMS資料とPOT資料による同一時系列資料の解析	75
6.2 形状母数が異なる原因および考えられる対応策	81
6.3 複数年で区切った最大値資料の解析	85
6.4 結論	92
参考文献	93
7. 結論	94
参考文献	98
Appendix	100
A-1 河川工事年表	101
A-2 建設省河川砂防技術基準（昭和33年）の序文	103
A-3 積率法とPWM法，L積率法	104
A-4 よく用いられる確率分布モデルのL積率解	110
A-5 一級河川における流域平均雨量のAMS資料の特性	120
参考文献	123

1. 緒論

自然災害は人類の社会的活動が自然現象による外力により被害を被ることであり、社会的活動の存在しないところで洪水氾濫等が起こってもそれは単なる自然現象であり災害ではない。

災害を引き起こした外力がその社会にとって対処可能な外力かどうか、また、次にやってくると想定される外力が技術的条件および社会経済条件の観点から対処しうるものかどうかなどが大きな関心事となる。

我々は大規模な自然現象や大災害に遭遇するとき、それに巻き込まれる程度により全く異なる反応をすることが多い。例えば、災害により大きな被害を受けた人からは、「また、次の台風で同じようなことが起きるかもしれない。どういう対策をしてくれるのか」と言われるが、一方で、被害を受けなかった人からは「こんな大災害はわれわれが生きているうちにはもう起きない」と言われる。これは、著者が2000年の東海豪雨の直後に矢作川中流部で実際に聞いた話である。後に統計の専門家から「天災は忘れた頃来る」^{注1}という名言で有名な寺田寅彦(1948)が「ものをこわがらな過ぎたり、こわがり過ぎたりするのはやさしいが、正当にこわがることはなかなかむつかしいことだと思われた」(昭和10年8月の浅間山の噴火の後の随筆)と書いていると教えていただき、まさにそのとおりになっていると実感した次第である。

このようなことがどの地域でも起きるとすると困ったことになる。すなわち、被災した人はまた被災するのが怖いから何とか対策をして欲しいと願っているし、一部の人が被災した状況を知っているが、自分が被災しなかった人は、何ら対策などする必要がないというように「対岸の火事」の認識である。このような状況は住民の意見を聞いて治水計画を立てる上で非常に大きな問題である。住民は各自が被災者という面で関係者でないかぎり対策に消極的な状況を維持し続けることになるので、多数決で決めると小規模な災害では後者の人が多い状況となり、大部分の人が被災しないかぎり抜本的な対策を行う計画に大多数の住民の賛同を得ることは一般に難しい。すなわち、このような状況が続けば、大多数の人が被災者となる大災害後に初めて認識することになる。住民は本当にそれを望んでいるのであろうか。ただ単にそのことに気づいていないだけなのではなかろうか。大多数の人が意見を表明しないことが多いわが国においては、まず、住民が正しく判断できる情報が住民にきちんと伝わり、かつ、常に未だ経験したことのないような災害に見舞われる可能性およびその場合の困窮度について他人の被災を「他山の石」として正しく認識し必要な対策を事前に行うことが重要である。

確率論的な現象を正しく理解することは一般に非常に難しい。世間では年超過確率1/50の現象が、50年に一度起きる現象と間違えて理解されることがしばしばである。正確にはある事象より大きな事象が起きる確率が1年間に1/50である事象をさして年超過確率1/50の事象と呼ぶ。いつ発生するかは全く分からないのであって、たまたま続いて発生することもある。非常に長い期間の観測結果からは概ね50年に1度の割合になるということであるが、この「年超過確率」の逆数を「再現期間」という専門用語で呼ぶことも誤解の遠因になっている。「年超過確率」より「再現期間」の方が確かに平易な言葉であるが、別の意味合いが入ってきていることが意外と認識されていない。このような状況では上述の大規模災害への正しい認識ができず、結局、災害の事前対策はいつまでもできないことになる。このため、事象を正當に評価し、住民に分かりやすく説明する専門家が必要となる。

注1 寺田寅彦の弟子である中谷宇吉郎は、昭和9年(1934)室戸台風の直後に「寺田が書いた「天災と国防」の中にこの言葉と同じ趣旨が書かれているが、このことと同じ表現はどこにも書かれていない。ペンを使わないで書いた言葉といえる」と述べている。なお、この名言は確率論からみても意味深長である。

わが国における自然災害のなかで最も頻度の高いもののひとつに洪水災害があげられるので、ここからは自然の外力が洪水の場合を考えてみる。すなわち治水対策において、計画目標としてどのような規模の外力を想定するかという問題である。平成9年(1997)に改正された河川法では、「河川管理者は、その管理する河川について、計画高水流量その他当該河川の河川工事及び河川の維持（次条において「河川の整備」という）についての基本となるべき方針に関する事項（以下「河川整備基本方針」という）を定めておかなければならない」（第十六条）とされている。より具体的には河川法施行令に河川整備基本方針に定める事項として、次のように定められており、洪水対策においては基本高水や計画高水流量、計画高水位が重要な要素として位置づけられている。

第十条の二 河川整備基本方針には、次に掲げる事項を定めなければならない。

一 当該水系に係る河川の総合的な保全と利用に関する基本方針

二 河川の整備の基本となるべき事項

イ 基本高水（洪水防御に関する計画の基本となる洪水をいう）並びにその河道及び洪水調節ダムへの配分に関する事項

ロ 主要な地点における計画高水流量に関する事項

ハ 主要な地点における計画高水位及び計画横断形に係る川幅に関する事項

ニ 主要な地点における流水の正常な機能を維持するため必要な流量に関する事項

河川法施行令第十条に、河川整備基本方針及び河川整備計画を作成する際、洪水、高潮等による災害の発生の防止又は軽減に関する事項については、過去の主要な洪水、高潮等及びこれらによる災害の発生の状況並びに災害の発生を防止すべき地域の気象、地形、地質、開発の状況等を総合的に考慮することと定められている。

また、国土交通大臣は、河川整備基本方針を定めようとするときは、あらかじめ社会資本整備審議会の意見を聴かなければならないこととされている。さらに、河川管理者は、河川整備計画の案を作成しようとする場合において必要があると認めるときは、河川に関し学識経験を有する者の意見を聴かなければならないことおよびこの場合に必要があると認めるときは、公聴会の開催等関係住民の意見を反映させるために必要な措置を講じなければならないことや関係都道府県知事又は関係市町村長の意見を聴かなければならないことが定められている。このように開かれた計画の作成過程においては計画目標は客観性・科学的合理性をもって決められるべきである。

わが国の治水対策は、明治以降、数年に1回程度の規模の実績洪水に対処してきた時代、既往最大洪水を対象としてきた時代を経て、より大きい未だ経験したことのない確率洪水を計画外力とするというふうに遷移してきている。既往最大の洪水への対策を行う段階までは観測記録をもとに対策を立てることができるが、既往最大より大きな外力を対象に計画を立てようとする場合、外力規模の設定方法が問題となる。米国では、当初、「最近 N 年間に観測された最大値の2倍の値」に耐えるような強いダムを建設すればよいとされていた。これは、事象の発生が指数分布にしたがう場合、観測期間内の最大値の2倍の大きさの事象の再現期間は観測期間の2乗に1よりも大きい因子をかけたものであることに基づいている。^{注2}（Gumbel 著 河田他訳、1963）

現在、わが国では計画規模の超過確率に対する事象（確率水文量）を求める際に主に毎年最大値を用いて頻度解析を行う方法が採用されている。

注2 1より大きい因子は指数分布の位置母数を尺度母数（標準偏差）で割った商の指数関数の値である。したがってこの値が大きい場合は観測期間の2乗よりかなり大きな再現期間になる。

本論文は比較的短期間の観測資料しかないわが国の水文資料蓄積状況を踏まえ、水文頻度解析を行う際の課題を見だしその解決方法について、実資料を用いて研究を行ったものである。

次章以降の本論文の構成は以下のとおりである。

第2章では、わが国における河川整備と水文頻度解析の歴史について述べ、既往最大洪水対応の河川整備から確率洪水規模を想定した河川整備への転換について述べるとともに水文頻度解析に用いることができる資料集積の現状と課題を述べる。

第3章では、適合度および安定性を考慮した水文頻度解析の一般手順および水文頻度解析によく用いられる確率分布モデルについて述べる。確率分布モデルの形状を規定する母数の推定法については、蓄積されている水文資料の観測期間が比較的短いことをふまえ、このような資料の解析に適している母数推定法について述べる。なお、平方根指数分布および平方根指数型最大値分布（江藤ら、1986）についてはその特性を詳しく述べる。また、本研究を行うにあたって筆者がMathematica^{注3}を用いて開発した水文極値の統計解析プログラムについて概要を紹介する。

第4章では、わが国の一級河川の資料を用いて年最大値系列を用いた水文頻度解析を行う。3時間、6時間、12時間、24時間の流域平均雨量資料から一般極値分布を当てはめた時の形状母数の地域変化をみるとともにL積率比(L-skewnessとL-kurtosis)の特性(Hosking and Wallis, 1997)について調べる。さらに洪水流量資料について適合度評価指標SLSC(Standard Least-Squares Criterion for goodness of fit)を用いた適合度（高棹ら、1986）について検討し、SLSCを用いる場合の基準について考察する。また、リサンプリング手法を用いた安定性評価を行いその有効性を検討する。

第5章では、非毎年値資料を用いた水文頻度解析について毎年値資料（年最大値系列）を用いた解析と比較する方法について述べるとともに、一級河川の3時間、6時間、12時間、24時間の流域平均雨量資料の非毎年値資料による解析と毎年値資料による解析結果を比較する。また、非毎年値資料を抽出する際の課題である閾値の選定方法についての検討を行うとともに適切な閾値を用いて抽出された非毎年値資料の解析結果と毎年値資料を用いた解析結果をGumbel確率紙上で比較し、問題点を抽出する。

第6章では、第5章で抽出された毎年値資料を用いた解析と非毎年値資料を用いた解析結果が異なることがあることに焦点をあて、その原因を調べる。従来、超過確率の小さい部分においては理論的にどちらの解析手法を用いても同じ結果を得ることができることから毎年値資料による解析が用いられてきたが両者の間に乖離がある場合があることおよび3母数の分布である一般極値分布や一般パレート分布を用いる場合その確率水文量の推定に形状母数が重要な役割を果たしており、形状母数を適切に設定する必要があることを示す。さらに、形状母数が異なる原因およびその場合に考えられる対応策について検討する。対応策の一つとして複数年を一区切りとした資料の抽出方法の有効性（田中、2004）について述べる。

第7章では、本論文で得られた成果をまとめるとともに今後に向けての課題などについて触れる。

注3 MathematicaはWolfram Researchが開発した技術・科学計算のための統合ソフトウェア環境で、数値計算、数式処理、視覚化、プログラミング機能を1つの柔軟でインタラクティブなソフトウェアシステムに統合している。

参考文献

- Gumbel, E. J., 河田龍夫, 岩井重久, 加瀬滋男監訳：グンベル 極値統計学, 生産技術センター新社, p.404, 1963.
- Hosking, J.R.M. and J.R. Wallis : Regional Frequency Analysis –An Approach Based on L-Moments–, Cambridge Univ. Press., p.207, 1997.
- 江藤剛治, 室田 明, 米谷恒春, 木下武雄：大雨の頻度, 土木学会論文集, 第369号/II-5, pp. 165-174, 1986.
- 高棹琢馬, 寶 馨, 清水 章：琵琶湖流域水文データの基礎的分析, 京都大学防災研究所年報, 第29号, B-2, pp.157-171, 1986.
- 田中茂信：複数年で区切った最大値資料の特性, 極値理論の工学への応用, 統計数理研究所共同研究レポート169, pp.1-7, 2004.
- 寺田寅彦：小爆発二件, 寺田寅彦随筆集, 第五卷, 岩波文庫, 岩波書店, 1948.
- 中谷宇吉郎：天災は忘れた頃来る, 中谷宇吉郎随筆集, 岩波文庫, 岩波書店, 1988.

2. 水文頻度解析と水工計画・河川整備

2.1 わが国の河川整備と水文頻度解析の歴史

わが国の河川整備の歴史について、真田(1956)は、「往古よりの治水事蹟を通観すれば、治水堤防は仁徳11年(323)当時の都たる難波の周辺大和川淀川の水交錯せる沮洳の地に新川を穿ちて排水を図り、又淀川に茨田堤を築きたるを史伝の最初とする」と述べており、Appendix A-1の年表に見られるように、近代以前からすでに多くの河川で堤防や灌漑施設等の整備が行われてきた。

明治になって欧米の技術がオランダ人技術者をはじめとする欧米の技術者の来日によりもたらされ、港湾の整備と一体となった河口部処理及び舟運を主目的とする低水工事が全額国費負担で国によって進められた。一方、高水工事は氾濫が一地域に限る地先の的なものとの認識から地方負担で整備された。明治18年の全国的な水害の発生により、利根川、信濃川、淀川、筑後川等で新たな計画の基に河川整備が実施されるようになった。続く明治20年代も大洪水が続き、高水工事についても国直轄工事が強く求められるようになり、ついに明治29年旧河川法が制定され、高水工事においても国直轄工事が可能になるとともに直轄工事以外の改良工事にも国庫補助する制度が定められた（松浦，1997，2000）。

その後、明治43年(1910)に全国的な大水害が発生し、勅令を以て内閣に臨時治水調査会が設置され、第一期治水計画の検討が行われ、国の事業として改修を進めるべき河川の選定、施工順位、工期及び工費が審議された。政府はこれを受け、財政的な裏付けをもって治水工事に着手した。大正10年(1921)には第二期治水計画が策定され、昭和8年(1933)には臨時治水調査会に代わり恒久機関として設置された土木会議によって第三期治水計画が策定された。第三期治水計画の中で中小河川を対象とした補助事業が組み込まれることになり、幹川から支川へ順次事業が展開された。このころの大規模工事として、荒川放水路や大河津分水路の建設等があげられるが、昭和初頭（1930年代中葉）までの大河川の改修によって、これまで氾濫を許容せざるを得なかった地域に対しても連続堤防が築かれ、河川の骨格が形成された。また、河川の工事は既往最大洪水を対象に進められたが、改修途中でその計画規模を超える洪水が発生した利根川等では、その度に計画の見直しを行った(大町,1999)。なお、計画の対象とする流量の生起確率としては、1934年の水害防止協議会^{注4}決定事項の中に、氾濫区域内の道路、鉄道などの路線のうち重要なものの高さは大体20年に1回起こる程度の洪水の水位（最大洪水位）より30cm以上を標準とするとしているので、この程度を洪水防御のための洪水規模と考えていたのであろう。ちなみに富永は計画の洪水流量としては、既往の最高流量を原則とするが、周囲の状況または財政の関係より10年に1回起こる程度の洪水流量で我慢することもあるとしている（山本，1999）。

1945年，1947年，1948年と1910年に匹敵する大災害が相次いで起こり，その対応として従来の河川改修計画を拡大した他に，できるだけ下流河川の拡幅を避けるため，貯水池，遊水地などが上流部に計画されている。また，多目的ダムの計画・着工が行われた。昭和28年(1953)は西日本大水害，台風13号による水害など中部以西に大水害が発生し，内閣に治山治水対策協議会が設置され，建設，農林両省にわたる治山治水事業の基本対策要項が決定されている。これは国営93河川，府県営1,142河川について洪水防御施設の整備を図ろうとするもので，内容としては河川改修とあわせて99カ所のダム，52万町歩の森林整備および大量の砂防工事が含まれている（山本，1958）。この要項の中で計

注4 室戸台風の大被害のあった1934年10月に設置された内務省21，農林省23，鉄道省22，通信省12，商工省3，宮内省1の合計82名の関係各省の技術官で構成される大協議会。

画論として注目すべきことは、「従来の計画洪水量および治水方式に再検討を加え、各水系毎に砂防施設、治水ダムおよび河川改修などの諸事業を総合した治水計画を樹立し、その積極的な遂行をはかる」ことである（山本，1993）。

一方，1951年頃から河川行政担当者の中で，計画手法の見直しを行っている。計画対象流量が何年に1回生じる流量であるのかを年超過確率で表し，全国の河川の間でバランスを取ろうとしている。この中で中安米蔵および米田正文の論文はその後の計画洪水流量に対する考え方の基礎を作った重要なものである。以下に「河川法全面改正に至る近代河川事業に関する歴史的研究」（山本，1993）および両者の論文から，関係部分の趣旨を抜粋する。

「治水計画における洪水流量について—千代川を中心として—」（中安，1950）によると，「満州事変以降大陸政策が強行されるに至って治水事業は昭和8年を頂点として低下の一途を辿り，第2次世界大戦に入ってからほとんど休止の状態となった。災害復旧費と改修工事費とのバランスが破れたのも昭和9年からである。現在の貧困な国家財政や災害で病弊した地方財政のもとでこの巨額の試算の支出を考える時，ここでは単純な技術的理想論は許されないであろう。今後の治水計画の基本方針は現実的，かつ科学的でなければならない。この観点より治水計画実施のあり方は従来治安の面を重視してあまり考慮されなかった経済的諸関係の調査の上に立たなければならない」とあり，経済的諸関係を明らかにした上での治水計画，治水事業の確立を強く主張している。計画洪水流量の確率年として，千代川は既定計画の17年に対し，43年が適当としている。他の河川では，天神川が既定計画50年に対し，46年，斐伊川下流が既定計画57年に対し55年，吉野川岩津下流が既定計画35年に対し93年と求めている。このように単に既定計画を上回る計画とすべきとしているのではなく，防災利益率を考慮して妥当な投資レベルという観点から見ている。なお，今後の課題として，(1)時間雨量の観測の充実，(2)指定水位以下の水位観測を流出の遅れ時間毎に実施すること，(3)ピーク付近の洪水観測は30分観測とすること，(4)全国的に雨量及び洪水観測資料の整理様式を統一すること，(5)水文調査のため，特定河川を指定して資料の充実を図ること，(6)水害の資料を充実することを強調している。

「淀川計画高水論」（米田，1952）によると，確率主義は，治水計画を理論的にし，安全，かつ信頼性のある計画と主張するが，既往最大洪水にも重きを置いている。また，計画高水とはそれまでのピーク流量のみの概念ではなく，ハイドログラフ（洪水波形）としており，この考えは後の「建設省河川砂防技術基準」における基本高水の概念と同様であり，この時代から本格的に建設されたダムの計画では必要不可欠であった。淀川の計画に対しては，「過去の洪水の水位および流量，特に既往最大流量および確率洪水流量を究明し，さらにユニットグラフにより降雨と流出の関係を明らかにし降雨記録より確率降雨を求め，それに対応する流量を求め，これらの条件により判断し既往最大流量を基本とし確実にして合理的な計画高水の最大流量を決定した」と述べている。

中安米蔵および米田正文の論文は昭和33年11月20日に発刊された「建設省河川砂防技術基準」に活かされている。

それでは，確率主義を実践するための方法論はどうであったかみてみる。中安(1950)は岩井(1947)の方法を用いており，米田(1952)は確率紙を用いて平分線を引く手法をとっている。河川砂防技術基準(1958)には，その解説に，「確率計算の方法としてはFuller, Hazen, Foster, Goodrich, Slade, Kimball, 岩井, Gumbel等の方法，非母数法と漸次改良され新しい提案がなされてきた。しかしいずれの方法とも全国一律の基準と決定するに至らないが，現在のところ岩井法を統計の要素の下限値が正又は零，上限値が無限大になるという条件のもとに採用するか，Gumbel法によるか，あるいは確率紙にHazenのプロットングポジションを用いてプロットし，平分線を引く簡略法を用いる場合が多い。ただし，簡略法は取り扱いに便利であるが，統計的健全性に欠け，客観性に乏しい欠点がある」という趣旨の説明がある。これと同時期に著された「河川工学」（山本，1958）には，Fullerの

方法、正規分布、対数正規分布、岩井法、Hazen紙（対数確率紙）による方法およびGumbelの方法が16頁にわたって説明されている。なお、Gumbelの方法については極値理論に基づくものと説明されている。したがって、このころ既に単に変量に対数変換して正規分布を適用するものと岩井法（対数正規分布）を用いる方法および極値分布（Gumbel分布）が使用されていたと考えられる。1958年に米国で出版されたGumbelの”Statistics of Extremes”の和訳「極値統計学」（Gumbel著、河田他訳、1963）が5年後に日本で出版された。この本の極値分布に関する章の和訳を担当している角屋は翌年「水文統計論」（角屋、1964）をまとめている。これによってわが国の水文頻度解析手法が体系的に一応確立され、この「水文統計論」と角屋の成果を多く取り入れた「応用水文統計」（岩井・石黒、1970）は実務家のバイブル的な役割を果たしながら現在に至っている（宝、1995）。その後、Greenwoodら(1979)により提案されたPWM（確率加重積率）法による正規分布、対数正規分布およびピアソンⅢ型分布の解が竹内・土屋(1988)により求められ、星(1998)は、その後の発展も含めてまとめている。なお、角屋の研究を調べてみると昭和28年(1953)にすでにGumbel分布について紹介しており（角屋、1953）、上述の和訳部分に関係が深い対数極値分布A型および対数極値分布B型について1956年に論文を発表しており、Gumbel分布と対数極値分布A型および対数極値分布B型が別々のものでなく歪み係数を通して一連の関係を有することを指摘しているとともにGumbel分布が対数正規分布のspecial caseであるとするVen Te Chow(1954)の論旨^{注5}に対し、批判している（角屋、1956）。残念ながら、いち早く発表されたJenkinson(1955)の一般極値分布(generalized extreme-value distribution)が今日世界的に有名になり、利用されている^{注6}。

次に、河川砂防技術基準の改訂経緯とその水文統計に関する記述についてみる。水文統計に関する部分についてみると、河川砂防技術基準は昭和33年の策定以来、昭和51年(1976)、52年(1977)、61年(1986)及び平成9年(1997)に改訂されている。

昭和33年河川砂防技術基準：基本高水のピーク流量と重要度との関係に関する記述において「基本高水のピーク流量の決定にあたっては、計画対象地域の重要度に応じて年超過確率を考慮するものとする。」とされており、解説において、確率計算の方法としてはFuller, Hazen, Foster, Goodrich, Slade, Kimball, 岩井, Gumbel等の方法、非母数法と漸次改良され新しい提案がなされてきた。しかしいずれの方法とも全国一律の基準と決定するに至らないが、現在のところ岩井法を統計の要素の下限値が正又は零、上限値が無限大になるという条件のもとに採用するか、Gumbel法によるか、あるいは確率紙にHazenのプロットングポジションを用いてプロットし、平分線を引く簡略法を用いる場合が多い。ただし、簡略法は取り扱いに便利であるが、統計的健全性に欠け、客観性に乏しい欠点があるという趣旨の説明がある。

昭和51年河川砂防技術基準（案）：主な改訂点として、最近実務上広く用いられ、また学術上の発展も著しい水文統計および流出計算に関する事項を基準化し、調査編にそれぞれ新しい章を設けたことがあげられる。第4章水文統計第3節水文学の生起確率の解析には38項を用いて、確率紙による簡略計算、分布関数式による確率計算およびその計算例を分かりやすく示している。取り上げている分布は

注5 石原・高瀬(1957)もVen Te Chow(1954)の理論の妥当性を積率理論によって根拠づけている。L積率比でも3母数対数正規分布とGumbel分布が近接することが分かる。しかし、Gumbel分布は一般極値分布の特殊形であり、3母数対数正規分布と一般極値分布の交点（共通点）がGumbel分布ではないので、角屋の指摘に理がある。

注6 角屋(1964)はJenkinson(1955)の方法とパラメータの変換で同等であることを示している。ただし、両者の方法およびL積率を用いた一般化極値分布では母数推定法が異なるので必ずしも全く同じ確率水文量を与えるとは限らない。

毎年値資料については、対数正規分布（岩井法、石原・高瀬法）、対数ピアソンⅢ型分布、Gumbel分布、対数極値分布（最大値分布および最小値分布）、非毎年値資料については指数分布が用いられている。

なお、改訂作業の過程における主要な検討事項として、「調査編の水文統計、流出計算、粗度係数および水位計算の各章に対しては、現在の実務担当者の仕事内容と比較して、本基準はやや学術的すぎるとの意見があった」と記している。

昭和52年改訂河川砂防技術基準（案）：河川管理施設等構造令及び同施行規則並びに公共測量作業規定が制定されたことに伴う一部改定が行われている。

昭和61年二訂河川砂防技術基準（案）：52年版と同一である。測量技術の向上等を背景に建設省公共測量作業規定が改正されたことに伴い、調査編の一部改訂を行ったものである。

平成9年改訂新版河川砂防技術基準（案）：改訂の背景として、新しい技術的知見を加えて修正（水文確率分布の適用分布形の選定方法を明記）したとされているが、第4章水文統計の第3節水流量の生起確率の解析にかかる部分は10頁弱と縮小されている。特に確率紙による簡略推定が残されているにもかかわらず、分布関数式による確率計算は3頁余りと大幅に縮小されており、具体的な計算方法等は文献を参照するほかはないにもかかわらず、参考文献のリストはあるが、具体的な対応が明確に示されておらず、実務的には不便なものになっている。平成17年に計画編が改訂され「国土交通省河川砂防技術基準」となっている。なお、一連の河川砂防技術基準（案）では極値分布の説明はあるものの極値の説明とその適用に関する説明は明確ではない。これは、変数を対数変換した値が正規分布に良く適合するとして経験的に対数正規分布が慣用されてきたことと無縁ではないと思われる。

実際の河川計画で確率に基づいた値が採用されたのは、昭和39年(1964)の石狩川が最初であった(山本, 1999)とされているが、「石狩川治水の曙光—岡崎文吉の足跡—」（北海道の治水技術研究会, 1990）によると「昭和28年9月神居古潭上流の区間の計画高水流量の検討の結果、確率評価1/100の計画降雨を用い、単位図法により伊納地点で $3,620\text{m}^3/\text{s}$ と定めた」とあり、これが最初の事例であると思われる。なお、この文献では昭和40年に石狩川の工事实施基本計画が改訂されており、「確率規模としての検討を全流域に対して行った初めてのものであり、水系を一貫して治水計画を策定したという意味で意義深いものである」とされている。この計画検討における確率評価の手法には岩井法と石原・高瀬法が用いられている。

米田(1952)が流域レベルで検討を行っていた淀川については、前述したように既往最大流量を基本とし「計画高水流量として80年洪水をとり $7,300\text{m}^3/\text{s}$ にすべきである」という意見書を上申したが、一方、京大の石原藤次郎博士は経済性の検討から淀川等の重要河川は80年洪水では不十分で数百年洪水をとっても経済的に過大といえないと論じた。このような理論的結論だけで計画高水流量を増加することは、事業規模が大きいだけに当時としては踏み切ることができず具体的な対策がまとまらないうちに昭和28年(1953)の13号台風の出水を迎えた」とあり、翌昭和29年河川審議会の慎重な審議を経て淀川水系改修計画が策定されている。この改修計画では台風13号出水の宇治川破堤がなかった場合の枚方における既往最大流量が大体1/100であることから、三支川の流量は13号出水の実績をそのままとらず、年超過確率1/80の計画とされている（淀川百年史, 1987）。このように、淀川水系改修計画では確率評価を行っているが、基本は既往最大流量に基づいている。河川法改正後昭和40年に策定された工事实施基本計画の基本高水は淀川水系改修計画の基本高水を用いており、年超過確率を1/200とした工事实施基本計画が定められたのは昭和46年であった。

極値に関する研究の歴史は古く、1709年Nicolaus Bernoulliの保険統計学の問題で最高生存者の平均寿命の研究が発端である。その後19世紀に、極値事象の発生回数をあらわすPoissonの研究および

その統計的意義と重要性を示したL. von Bortkiewiczの研究を経て、1920年代のGumbel, R. von MisesおよびL.H.C. Tippettの研究につながっている。最初は、正規分布から抽出された標本の最大値について研究されたが、このことがその理論の発展を妨げたようであり、後に元の分布が正規分布と異なる概念の研究をM. Fréchet(1927)が発表し、翌年にR.A. Fisher and L.H.C. Tippett(1928)が極値理論に関する研究をまとめた（Gumbel著、河田他訳、1962）。これにGnedenko(1943)の研究が加わり、極値分布には3つの型があつて、それ以外にないことやそれらの極値分布の吸引領域に分布が属する^{注7}ための必要十分条件が明らかになった。これらの3つの型の極値分布を一つの式で表したものが前述のように一般極値分布と呼ばれ、区分最大値（1年間等のある区分内のデータの最大値）のデータ解析に応用されている。区分最大値のみでは利用可能データの個数は限られて精度のよい解析が行われない場合が多いが、水文学の分野で使われていたPOT(Peaks Over Threshold)法は十分大きな閾値を超過するデータが指数分布で近似できるという事実に基づく解析手法であり、これに関連した一般パレート分布がBalkema and de Haan(1974)とPikands(1975)により導出されている。一般極値分布と一般パレート分布の形状母数は相互に共通の形状母数を有している（高橋、2008）。

一方、わが国の水文統計解析では、対数正規分布が重用され、いろいろな改良が行われたが、極値分布に関する検討は角屋の一連の研究を除いてあまり行われてこなかった。また、角屋(1961)は年最大値(AMS; Annual Maximum Series)資料と非毎年値(POTまたはPDS; Partial Duration Series)資料の理論的関係や応用例を示す中で、資料整理の簡単なAMS解析を行うことにより、T年に1回起こるような事象から1年に何度も起こるような事象まで推定することが可能であり、AMS資料かPOT資料という問題に対してはAMS資料で実用上十分であるとしており、AMS解析への偏重に拍車をかけたきらいがある。なお、角屋のこの研究は地点雨量記録を対象に行われている。著者らは洪水ピーク流量を対象に解析を行い、AMS資料とPOT資料が必ずしも同じ確率水文量を与えるとは限らないことおよびこの原因としてAMS資料における課題を見いだした（田中・宝、2001）。宝・高棹(1988)は分布モデルの評価基準として適合度のみでなく確率水文量の安定性も考慮すべきことを示している。さらに、宝・高棹(1990)は3母数対数正規分布、Gumbel分布一般極値分布の母数推定法の比較評価を行い小標本の場合L積率法が優れている推定法であることを示している。これらの研究の背景には計算機の発達や新しい統計技術がある。近年の情報機器の発達により、観測や資料整理がずいぶん楽になってきており、非毎年値資料としての時系列資料の整理も十分行える時代になってきている。AMS資料に比べPOT資料は多くのデータを扱うことができるので、気候変動の影響等の解析にも有利である。本論文ではこのような時代に必要となる解析法についてわが国の直轄河川の資料を基に検討する。

2.2 わが国の水文資料の現状と課題

河川計画の基本となる主な外力は、対象とする施設により水位であつたり、ピーク流量や流量ハイドログラフのボリュームであつたりするが、基本となるのはピーク流量を含む流量ハイドログラフである。わが国の直轄河川の流出は日のオーダーの現象であり、その変化をみるには時間雨量や毎時の流量が基本量として必要である。流域が比較的小さい場合には時刻流量だけでなくピーク流量も重要となるので一般に時刻流量の最大値とは別にピーク流量もあわせて整理蓄積されており、流量年表に

注7 母集団が指数分布、ガンマ分布、正規分布、対数正規分布、ワイブル分布、一様分布など教科書に出てくるほとんどの分布は極値分布で近似できるという意味である。Gumbel分布の吸引領域に含まれるものに、指数分布、ガンマ分布、正規分布、対数正規分布、ロジスティック分布、双曲分布、一般双曲分布、負のフレシェ分布がある（志村、2008）。

はこの年最大値も載っている。なお、直轄河川の重要な区間はほとんど堤防で守られており、その堤防の安定を考える上ではピーク水位だけでなく、雨量およびそれぞれの継続時間が重要となる。また、流量は流域の変化の影響を受けやすいが、降雨は自然現象であり人為的な影響を受けないと考えられること、さらに守られる地域の安全性を考える際、外水に対する安全度だけでなく内水氾濫についても考える必要があることから降雨量の資料は最も重要な資料である。

河川統制事業調査が昭和12年(1937)に開始されており、雨量、水位、流量の観測が開始されており、他に地形地質調査、利用状況調査等から構成されている。本調査は内務省、通信省、農林省で分担され、第1期調査分として63河川が取り上げられ、内務省はそのうち37河川を担当している。なお、利根川では大正4年から流量観測が行われており、一部の河川では河川統制事業調査よりも前から流量観測が行われている(山本, 1993)。

1999年時点で、各地方建設局から提供いただいた資料によると、直轄河川に関する時間雨量などの比較的詳しい水文資料の多くは昭和30年前後から整備され始めており、一部には戦時中を通して観測が継続されていた観測所もあった。図2.1は1999年時点で、整理されていた全国の一級水系における137カ所の年最大流量の連続した観測年数を示したものである。40年～44年が48標本、35年～39年が29標本、45年～49年が28標本、50年～54年が9標本であり、平均値は44.1年である。現在ではこれらのデータは単純に10年蓄積が増えていると考えられる。

例えば、東海地方を流れる豊川については、水位流量観測は大正8年(1919)に開始されている。降雨量については日雨量が明治26年(1893)と早くから観測されており、観測期間が100年を超えるが、時間雨量の観測資料は昭和30年(1955)からである。

観測流量は観測地点より上流の氾濫やダム等による貯留の影響を受けている可能性がある。このような流域について、河川計画の基本となる基本高水流量を流量資料から検討する場合には、これらの影響を流出解析により取り除いた流量を用いる必要がある。その影響を取り除くために時間雨量観測データが必要となる。一般に、河川計画においては計画降雨継続時間を設定し、その時間内に降った降雨量を統計解析の対象資料とする。時間雨量の観測が開始されるまでは日雨量のみしか観測されておらず、上述の補正を行うことはできないが、計画降雨継続時間が1日や2日の場合、日降雨量や2日雨量を確率評価することは可能である。この場合、計画規模のハイドログラフを得るにはいくつかの既往の大洪水の降雨パターンをあてはめて時間雨量(ハイトグラフ)を作成するのが一般的である。

一方、近年は温暖化ガスの排出による地球温暖化およびそれに伴う気候変化の進行がほぼ確実となり、その影響が懸念されている。降雨量も地球温暖化の影響を受けて変化しており、今後ますます大きな変化が起きると想定されている(Meehl, et al., 2007)。これまで水文頻度解析においては時系列

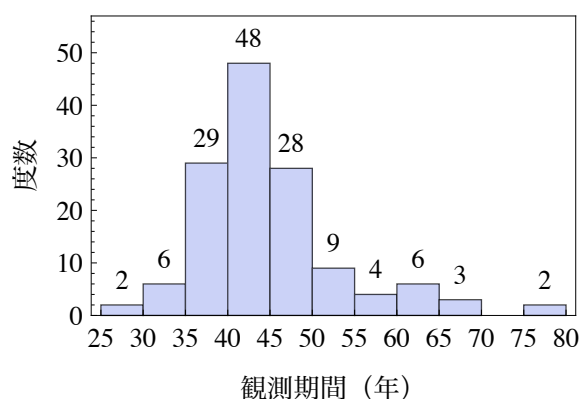


図2.1 連続した流量観測資料の期間のヒストグラム

資料から年最大値資料を抽出し、それを解析対象としてきたが、より多くの資料を対象に解析できるような体制を整備することが重要である。

2.3 結論

わが国の河川整備の歴史と水文頻度解析の関係について調べてみた。戦後の災害が頻発する中で、災害発生の有無がその河川整備の水準にも影響することおよび一度既往の出水規模を上回るような災害に見舞われるとバランスのとれた河川整備を差し置いて大規模な整備が始まること等の問題点も多い。単に災害復旧を実施するだけでなく、バランスのとれた河川整備は如何にあるべきかを考えた結果が確率主義の導入であった。これは河川整備が防災便益とバランスすべきという考えであり、このためには精度よく外力の確率評価と防災便益の評価を行う必要がある。

本来、各河川の洪水到達時間に基づいて決められる計画降雨継続時間内の降雨資料を整理し、その年最大値資料から確率水文量が求められるべきであるが、確率主義の黎明期には、それまでのデータの蓄積が少ないことから、比較的観測期間の長い日雨量や2日雨量および3日雨量が使われることが多かった。このため、日雨量や2日雨量および3日雨量の年最大値を統計解析の対象とした。確率雨量が求められた後は、降雨波形を既往の洪水時の波形等を基に設定し、単位図法等により流出計算してピーク流量や流出ボリュームを算定している。米田は流出計算に信頼が置けなかったようで、既往最大流量を基本とした考え方をとっているし、河川砂防技術基準もいろいろな要素について確率評価をするよう求めている。

しかしながら、河川砂防技術基準策定当時建設技監であった米田は、「技術は段階的に進歩するものである。この段階を示すのがこの技術基準である。すなわち、今日の段階における建設省の河川に関する技術水準はこの基準書に示す通りであるが、来年の技術水準はこの基準書よりもさらに進歩すべきものであって、できれば毎年この基準書は改訂を加えてゆくべき性格のものである。かようにすることによって建設省の技術基準は毎年進歩し、しかもその進歩の跡が明らかになり、新進者の研究を促進することになる。かくて河川に関する技術は秩序よく、能率的に、加速度的に前進するであろうと思う。これが私のこの技術基準書を制定しようとした動機であるが、この基準が将来の河川技術の基礎となってけんらんたる技術の殿堂が築かれることを固く信じている。」とその序文^{注8}に著している。この認識はニュートンが遺した「私がほかの人より遠くを見ることが可能だったのは、私が巨人の肩に立ったからである」という一節にみられる先人の業績がなければこれだけの成果を挙げられなかっただろう（黒田，2002）という意味の言葉に通じるところがあり、常に技術の改良を積み重ねていかねばならない。一方で、「巨人の肩に立つ」までが大変であることも事実であるので、技術水準が分かりやすく体系化されて広く普及されることも重要である。

水文極値の統計解析は河川計画を検討するにあたり基本的なものであり、客観的合理性を有する河川整備基本方針や河川整備計画を検討するためには必要不可欠な技術である。今後、気候変動の影響等これまで考えてこなかった要素も考慮しなければならなくなる。快適かつ便利になった現在の社会を支えているものの大部分は電気電子製品であり水に弱い。2005年に発生した米国ニューオーリンズのハリケーンカトリーナによる大災害等大規模な水災害は現代社会の活動継続に対して時間的にも空間的にも大きな影響を与える。世界第2位の経済大国であるわが国がおかれている自然条件は元来厳しい。守るべき社会的な資産は非常に大きくなっているにもかかわらず、治水計画の計画規模は昭和33年当時の2倍程度にしか設定されていない。資産が増えれば大きな保険をかけたり防災システムを

注8 Appendix A-2にその全文を示す。なお、「まえがき」は河川局長山本三郎、「基準書作成の経過について」は河川計画課長中安米蔵が執筆している。

より強固にするのは当然であろう。気候変動に適応した治水対策検討小委員会(2007)によると、オランダでは1985年には1/10,000規模の安全度の高潮防御システムが整備され、イギリスのテムズ川には1983年に1/1,000の整備が完成している。国際社会で協調していくためには万が一の時に支援し合うのも重要であるが、普段からの備えも同程度のものが要求されていると考えるべきである。直轄河川の水文資料蓄積状況からすると、確率規模が1/100と小さければデータの蓄積によりそのうち大きな外挿をしなくてよい時代が来るが、諸外国のように大きな確率規模に対する確率水文量を推定するには少々のデータの蓄積では外挿しなくて済むようにならない。したがって、水文資料の蓄積を継続するとともに極値資料の解析技術についても検討を積み重ねる必要がある。

参考文献

- Chow, V.T. : The Log-probability Law and its Engineering Applications, Proc. A.S.C.E., Vol.80, No.536, 1954.
- Fisher, R.A. and L.H.C. Tippett: Limiting forms of the frequency distribution of the largest or smallest member of a sample, proc. Cambridge Philos. Soc., Vol.24, p.180-190, 1928.
- Fréchet, M.: Sur la loi de probabilité de l'écart maximum, Annales de Société Polonaise Mathématique, Vol.6, pp.92-116, 1927.
- Greenwood, J.A., J.M. Landwehr, N.C. Matalas and J.R. Wallis: Probability Weighted Moments—Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form, Water Resources Research, Vol. 15, No.5, pp.1049-1054, 1979.
- Gumbel, E. J., 河田龍夫, 岩井重久, 加瀬滋男監訳: グンベル 極値統計学, 生産技術センター新社, p.404, 1963.
- Jenkinson, A.F.: The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements, Quarterly Journal of the Royal Meteorological Society, 81, pp. 158-171, 1955.
- Meehl, G.A., T.F. Stocker, W.D. Collins, P. Friedlingstein, A.T. Gaye, J.M. Gregory, A. Kitoh, R. Knutti, J.M. Murphy, A. Noda, S.C.B. Raper, I.G. Watterson, A.J. Weaver and Z.-C. Zhao, 2007: Global Climate Projections. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Pickands, J.: Statistical inference using extreme order statistics. Ann. Statist. 3, pp.119-131, 1975.
- 石原藤次郎, 高瀬信忠: 対数正規分布とその積率による解法, 土木学会論文集, 第47号, pp.18-23, 1957.
- 岩井重久: 水文学における非対称分布に就いて, 土木学会論文集, 第1, 2号合併号, pp.93-116, 1947.
- 岩井重久, 石黒政儀: 応用水文統計学, 森北出版, p.370, 1970.
- 大町利勝: THE RIVER LAW WITH COMMENTARY BY ARTICLE—Legal Framework for River and Water Resources Management in Japan—, IDI Water Series No. 4, Infrastructure Development Institute, 1999.
- 角屋 睦: 計画洪水量に関する順序統計学的考察, 農業土木研究, 第21巻, 第3号, pp.188-195, 1953.
- 角屋 睦: 極値分布とその一解法, 農業土木研究, 第23巻, 第6号, pp.350-357, 1956.
- 角屋 睦: 雨量分布とその年最大値の分布, 京大防災研究所年報, 第4号, pp.122-131, 1961.
- 角屋 睦: 水文統計論, 水工学シリーズ, 土木学会水理委員会, 64-02, p.59, 1964.
- 気候変動に適應した治水対策検討小委員会: 第1回小委員会資料, 資料3, 2007
- 黒田玲子: 科学を育む, 中公新書, 中央公論新社, p.256, 2002.

建設省河川局編纂：建設省河川砂防技術基準，1958.

建設省河川局監修：建設省河川砂防技術基準（案），調査編，1976.

建設省河川局監修：改訂建設省河川砂防技術基準（案），調査編，1977.

建設省河川局監修：二訂建設省河川砂防技術基準（案），調査編，1986.

建設省河川局監修：改訂新版河川砂防技術基準（案）同解説，調査編，1997.

真田秀吉：明治前日本土木史，日本学術振興会，p.554，1956.

志村隆彰：吸引領域と離散分布，極値理論の工学への応用(5)，統計数理研究所共同研究リポート 212，統計数理研究所，pp.25-32，2008.

高橋倫也：極値統計学，統計数理研究所，公開講座資料，p.57，2008.

宝 馨：水文頻度解析の古くて新しいパラダイム，土木学会第50回年次学術講演会，1995.

宝 馨，高棹琢馬：水文頻度解析における確率分布モデルの評価規準，土木学会論文集，第393号/II-9，pp.151-160，1988.

宝 馨，高棹琢馬：水文頻度解析モデルの母数推定法の比較評価，水工学論文集，土木学会，第34巻，pp.7-12，1990.

竹内邦良，土屋一仁：正規分布，対数正規分布およびピアソンⅢ型分布のPWM解，土木学会論文集，第393号/II-9，pp.95-101，1988.

田中茂信，宝 馨：洪水頻度解析におけるAMSとPDSの比較，水工学論文集，第45巻，土木学会水理委員会，pp.205-210，2001.

中安米蔵：治水計画における洪水流量について—千代川を中心として—，p.145，1950.

北海道の治水技術研究会：石狩川治水の曙光—岡崎文吉の足跡—，北海道開発局，p.261，1990.

松浦茂樹：国土づくりの礎 川が語る日本の歴史，鹿島出版会，p.216，1997.

松浦茂樹：戦前の国土整備政策 日本経済評論社，p.353，2000.

山本晃一：河道計画の技術史，山海堂，p.658，1999.

山本三郎：河川工学，朝倉書店，p.518，1958.

山本三郎：河川法全面改正に至る近代河川事業に関する歴史的研究，（財）国土開発技術研究センター，p.461，1993.

淀川百年史編集委員会：淀川百年史，建設省近畿地方建設局，1987.

米田正文：淀川計画高水論，近畿地方建設局，p.161，1952.

3. 水文頻度解析の手法

3.1 頻度解析の一般的手順

水文統計解析でよく行われることは、(1)標本が属する母集団の特性を推定すること、(2)確率分布モデルを用いて任意の超過確率に対する確率水文量を推定すること、(3)確率変量の値が与えられたときにその値の起こりやすさ(超過確率や再現期間で表すことが多い)を求めること、(4)標本に適合する確率分布モデルを見いだすこと、(5)標本と確率分布モデルを確率紙に描画すること、(6)標本のヒストグラムと確率分布モデルの確率密度関数を比較すること、(7)安定性のよい確率分布モデルを見いだすことなどである。母集団の特性は標本を通してしか推定することができず、上述の作業を行うためには一般的に確率分布モデルを選択し、適切にその母数を推定する必要がある。

確率統計解析を行う際の前提条件は、標本は母集団からある一定の条件を満たすように抽出されたものであることである。その条件とはランダム性、独立性、均質性、斉次性を有することである。

1) ランダム性(Randomness)

水文学的には自然の原因による変動があること、人為的に調節された流量データはランダムとは見なされないので、調節効果を除去する必要がある。

2) 独立性(Independence)

標本データがそれぞれ独立であること。あるデータが他のデータの影響を受けていないこと。例えば、本日の流量は昨日の流量の影響を受けている（昨日の流量が大きければ、本日の流量も大きい）ので独立とは言えない。通常、データのサンプリング間隔を大きくすれば、独立と見なせる。

3) 均質性(Homogeneity)

ひとつの母集団からの標本と見なせること。融雪出水と降雨による出水のデータは均質とはみなせないかもしれないが、検定テストの結果により均質とみなせるかもしれない。

4) 斉次性(Stationarity)

ランダムな変動を除いた後の成分が時間的に変化していないこと。非斉次なデータは、トレンド、ジャンプ、周期をもつ。例えば、トレンドは徐々に進む流域の改変、ジャンプはダム建設、周期は気候変動と関係する。

しかしながら、確率統計解析の対象となる水文資料には種々の問題が含まれており、確率統計解析にあたってはデータの独立性・等質性を仮定せざるを得ないという状況にある（寶，1989）。

一昔前までは、水文頻度解析に図解法が用いられていたが、計算機の発達にあわせていろいろな解析手法が開発され、実務においては、今では標本の統計量から確率分布モデルの母数（パラメータ）を推定する手法が水文頻度解析における一般的手法となっている。

確率紙による図解法では判断が主観的であるという大きな問題があったが、確率分布モデルを用いる方法に変わったからといって問題がなくなったわけではない。治水計画に関係する頻度解析においては主に極値資料が用いられるが、これを対象とする確率分布モデルには極値理論に基づくもの及び従来から慣用的に用いられてきたもの等いくつものモデルが存在し、かつ、それぞれの母数推定法もいくつも提案されている。このような状況の中でどのように客観的合理性を持たせて適切なモデルおよび母数推定法を選択するかということが問題となってきた。

このような状況をふまえ、宝・高棹(1988)は水文頻度解析による現代的な確率水文量の推定手法として「データに対する適合度が良く、かつ、確率水文量の変動が小さいモデルが実用上望ましい」という観点から、次のように提案している。

- Step 1) データの吟味
- Step 2) 候補モデルの列挙
- Step 3) 母数推定
- Step 4) 候補モデルのscreening
- Step 5) 確率水文量の変動性の検討
- Step 6) 最終モデルの決定

この手順にしたがって、各段階の検討を次のように行う。

- Step 1) データの吟味：観測された値に誤記などのミスがないかどうか細心のチェックを行うことはもとより、確率統計解析を行う際の前提条件であるデータのランダム性・独立性・均質性・斉次性の観点から検討する（前述）。
- Step 2) 候補モデルの列挙：ヒストグラムや分布曲線を描き大体の分布形状を把握した後、水文統計でよく使われる確率分布モデルの中から標本の大きさを考慮して使用するモデルを列挙する。
- Step 3) 母数推定：確率分布モデルの母数を標本の統計量を用いて求める。方法としては最尤法、積率法、PWM法、L積率法などがある。標本サイズによりバイアスや誤差が異なるので適切な手法を用いる必要がある。
- Step 4) 候補モデルのscreening：客観的な適合度指標を用いて適合度の悪いモデルを候補モデルから除外する。
- Step 5) 確率水文量の変動性の検討：Step 4) で残ったモデルについて、リサンプリング手法を用いて確率水文量の変動性を調べる。
- Step 6) 最終モデルの決定：確率水文量の変動が最も小さいモデルを選択する。

3.2 水文頻度解析によく用いられる確率分布モデル

極値水文資料を対象にする水文頻度解析において、多くの確率分布モデルが用いられてきている。複数のモデルについて具体的な解説が行われているものとして、初期には角屋(1964)による「水文統計論」、岩井・石黒(1970)の「応用水文統計学」があげられる。近年では、水文・水資源ハンドブック(1997)の「7.3 水文頻度解析」がかなり詳しく紹介している。また、海外では、Handbook of Hydrology(1993)の第18章にStedingerらによる“Frequency Analysis of Extreme Events”があり、水文極値解析に関する幅広い解説が述べられている。これらの文献で扱われている分布モデルと母数推定法について表-3.1に示す。

この表をみると「水文頻度解析」およびHandbook of HydrologyにおいてPWM法およびL積率法が加わっていることがわかる。PWM(Probability Weighted Moments)法はGreenwoodら(1979)により提案された方法であり、L積率法(L-Moments)はそれを直感的に分かりやすく書き換えたものである。また、Regional Frequency Analysis(Hosking and Wallis, 1997)はL積率に関する詳しい解説と各分布のL積率による母数推定法を示している。詳細はAppendix A-3およびA-4に紹介する。

表3.1 水文頻度解析によく用いられる確率分布モデル

文献名 分布モデル	母数 の数	水文統計論 (1964)	応用水文統計学 (1970)	水文頻度解析 (1997)	HANDBOOK OF HYDROLOGY (1993)	REGIONAL FREQUENCY ANALYSIS (1997)
一様分布	2				L積率法	L積率法
指数分布	2	積率法	積率法		積率法、L積率法	L積率法
一般パレート分布(GP)	3				積率法、L積率法	L積率法
双曲線指数分布	3	積率法				
Gumbel分布	2	積率法	積率法	積率法、L積率法	積率法、L積率法	L積率法
対数極値分布A型	3	積率法	積率法			
対数極値分布B型	3	積率法 ¹⁾	積率法 ¹⁾			
一般極値分布 (GEV)	3	Jenkinsonの方法 (積率法)		積率法 ²⁾ 、L積率 法	積率法 ²⁾ 、L積率 法	L積率法
Weibull分布	3	積率法	積率法			L積率法
正規分布	2			積率法、PWM法	積率法、L積率法	L積率法
対数正規分布	3	石原・高瀬の方法 ³⁾ 、岩井法、順序確 率法	石原・高瀬の方法 ³⁾ 、岩井法、順序確 率法	積率法、クオンタ イル法、PWM法	積率法、L積率法	L積率法
ピアソンⅢ型分布	3			積率法、PWM法	積率法	積率法、L積率 法
対数ピアソンⅢ型分布	3			対数標本積率 法、対数標本 PWM法、原標本 積率法	積率法	
平方根指数型最大値分 布 (SQRT-ETmax)	2			最尤法 ⁴⁾		
一般ロジスティック分 布	3					L積率法
Kappa分布	4					L積率法
Wakeby分布	5					L積率法

注1) 水文統計論および応用水文統計学では対数極値分布B型の場合は危険側になるのを避けるため、Gumbel分布を当てはめるべきであるとして、母数の解法は詳しく示されていない。

注2) GEV分布については積率と母数の関係式は示されているがその具体的な解法は示されていない。

注3) 石原・高瀬の方法は積率法的一种である。

注4) 最尤法の解を繰り返し計算で求める方法が示されている。

水文頻度解析においてよく用いられてきている1変数の確率分布モデルには次のようなものがある。

- ・正規分布系のモデル：正規分布，対数正規分布
- ・ガンマ分布系のモデル：指数分布，ピアソンⅢ型分布
- ・極値分布系のモデルおよび関係する分布モデル：グンベル分布，一般極値分布およびこれらに関係が深い指数分布，一般パレート分布，ロジスティック分布，一般ロジスティック分布
- ・平方根指数分布系のモデル：平方根指数分布，平方根指数型最大値分布

以下にこれらの確率分布モデルの密度関数と分布関数を示す。

1) 正規分布系のモデル

【正規分布】

正規分布は水文極値資料の頻度解析において必ずしもよく使われるものではないが，他の分布の形状母数が特殊な場合に正規分布になる。

2母数：位置 μ ，尺度 σ ， x の分布範囲： $-\infty \leq x \leq \infty$

$$\text{確率密度関数：} f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} = \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \quad (3.1)$$

$$\text{累積分布関数：} F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left\{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right\} dt = \Phi\left(\frac{x-\mu}{\sigma}\right) \quad (3.2)$$

$$\text{ここに，} \phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right), \quad \Phi(x) = \int_{-\infty}^x \phi(t) dt \quad (3.3)$$

【対数正規分布】

3母数：位置 ξ ，尺度 α ，形状 k ，（ $k=0$ の場合は正規分布）

$$x\text{の分布範囲：} \begin{cases} -\infty < x \leq \xi + \alpha/k & (k > 0) \\ -\infty < x < \infty & (k = 0) \\ \xi + \alpha/k \leq x < \infty & (k < 0) \end{cases}$$

$$\text{確率密度関数：} f(x) = \frac{\exp\{k y - y^2/2\}}{\alpha \sqrt{2\pi}}, \quad y = \begin{cases} -k^{-1} \log\{1 - k(x - \xi)/\alpha\} & k \neq 0 \\ (x - \xi)/\alpha & k = 0 \end{cases} \quad (3.4)$$

$$\text{累積分布関数：} F(x) = \Phi(y) \quad (3.5)$$

ここに， Φ は標準正規分布の分布関数であり，式(3.3)で表される。

2) ガンマ分布系のモデル

【指数分布】

2母数：分布の下限 ξ ，尺度 α ， x の分布範囲： $\xi \leq x \leq \infty$

$$\text{確率密度関数：} f(x) = \frac{\exp\left\{-\frac{x-\xi}{\alpha}\right\}}{\alpha} \quad (3.6)$$

$$\text{累積分布関数：} F(x) = 1 - \exp\left\{-\frac{x-\xi}{\alpha}\right\} \quad (3.7)$$

【ピアソンⅢ型分布】

3母数：位置 μ , 尺度 σ , 形状 γ

形状母数 γ が0でなければ次の変数変換をおこなうと, γ の正負毎に x の分布範囲, 確率密度関数, 累積分布関数が以下のように表される.

$$\alpha = \frac{4}{\gamma^2}, \quad \beta = \frac{1}{2} \sigma |\gamma|, \quad \xi = \mu - 2 \sigma / \gamma, \quad \gamma \neq 0 \quad (3.8)$$

$\gamma > 0$ の場合, x の分布範囲: $\xi \leq x < \infty$

$$\text{確率密度関数: } f(x) = \frac{(x - \xi)^{\alpha-1} \exp\{-(x - \xi)/\beta\}}{\beta^\alpha \Gamma(\alpha)} \quad (3.9)$$

$$\text{累積分布関数: } F(x) = G\left(\alpha, \frac{x - \xi}{\beta}\right) / \Gamma(\alpha) \quad (3.10)$$

$\gamma < 0$ の場合, x の分布範囲: $-\infty < x < \xi$

$$\text{確率密度関数: } f(x) = \frac{(\xi - x)^{\alpha-1} \exp\{-(\xi - x)/\beta\}}{\beta^\alpha \Gamma(\alpha)} \quad (3.11)$$

$$\text{累積分布関数: } F(x) = 1 - G\left(\alpha, \frac{\xi - x}{\beta}\right) / \Gamma(\alpha) \quad (3.12)$$

ここに, $G(\cdot)$ は不完全ガンマ関数

$$G(\alpha, x) = \int_0^x t^{\alpha-1} \exp(-t) dt \quad (3.13)$$

である.

Pearson III型分布は形状母数 $\gamma=2$ の場合, 指数分布, $\gamma=0$ の場合, 正規分布, $\gamma=-2$ の場合, 逆指数分布となる.

3) 極値分布系のモデルおよび関係する分布モデル

【ゲンベル分布】

2母数：位置 ξ , 尺度 α , x の分布範囲: $-\infty \leq x \leq \infty$

$$\text{確率密度関数: } f(x) = \frac{\exp\left\{-\frac{x-\xi}{\alpha}\right\}}{\alpha} \exp\left[-\exp\left\{-\frac{x-\xi}{\alpha}\right\}\right] \quad (3.14)$$

$$\text{累積分布関数: } F(x) = \exp\left[-\exp\left\{-\frac{x-\xi}{\alpha}\right\}\right] \quad (3.15)$$

【一般極値分布】

3母数：位置 ξ , 尺度 α , 形状 k

$$x \text{の分布範囲: } \begin{cases} -\infty < x \leq \xi + \alpha/k & (k > 0) \\ -\infty < x < \infty & (k = 0) \\ \xi + \alpha/k \leq x < \infty & (k < 0) \end{cases}$$

$$\begin{aligned}
\text{確率密度関数} : f(x) &= \frac{1}{\alpha} \left(1 - \frac{k(x-\xi)}{\alpha} \right)^{\frac{1}{k}-1} \exp \left\{ - \left(1 - \frac{k(x-\xi)}{\alpha} \right)^{1/k} \right\} \\
f(x) &= \frac{1}{\alpha} \exp \{ -(1-k)y - \exp(-y) \}, \\
y &= \begin{cases} -k^{-1} \log \{ 1 - k(x-\xi)/\alpha \} & k \neq 0 \\ (x-\xi)/\alpha & k = 0 \end{cases}
\end{aligned} \tag{3.16}$$

$$\text{累積分布関数} : F(x) = \exp \left[- \left(1 - k \frac{x-\xi}{\alpha} \right)^{1/k} \right] = \exp \{ -\exp(-y) \} \tag{3.17}$$

以下に極値分布と関係の深いモデルを示す。

【指数分布】

上述、次に示す一般パレート分布の形状母数 $k=0$ の場合でもある。

【一般パレート分布】

3 母数：位置 ξ , 尺度 α , 形状 k , x の分布範囲 : $\xi \leq x \leq \xi + \alpha/k$ ($k > 0$), $\xi \leq x < \infty$ ($k \leq 0$)

$$\begin{aligned}
\text{確率密度関数} : f(x) &= \frac{1}{\alpha} \left\{ 1 - k \frac{x-\xi}{\alpha} \right\}^{1/k-1} = \frac{1}{\alpha} \exp \{ -(1-k)y \}, \\
y &= \begin{cases} -k^{-1} \log \{ 1 - k(x-\xi)/\alpha \} & k \neq 0 \\ (x-\xi)/\alpha & k = 0 \end{cases}
\end{aligned} \tag{3.18}$$

$$\text{累積分布関数} : F(x) = 1 - \left(1 - k \frac{x-\xi}{\alpha} \right)^{1/k} = 1 - \exp(-y) \tag{3.19}$$

【ロジスティック分布, 一般ロジスティック分布】^{注9}

3 母数：位置 ξ , 尺度 α , 形状 k , ($k=0$ の場合は logistic 分布)

$$x \text{ の分布範囲} : \begin{cases} -\infty < x \leq \xi + \alpha/k & (k > 0) \\ -\infty < x < \infty & (k = 0) \\ \xi + \alpha/k \leq x < \infty & (k < 0) \end{cases}$$

$$\begin{aligned}
\text{確率密度関数} : f(x) &= \frac{\alpha^{-1} \exp \{ -(1-k)y \}}{(1 + \exp(-y))^2}, \\
y &= \begin{cases} -k^{-1} \log \{ 1 - k(x-\xi)/\alpha \} & k \neq 0 \\ (x-\xi)/\alpha & k = 0 \end{cases}
\end{aligned} \tag{3.20}$$

$$\text{累積分布関数} : F(x) = 1 / \{ 1 + \exp(-y) \} \tag{3.21}$$

図3.1は種々の分布のL-skewness τ_3 とL-kurtosis τ_4 を比較したものである。2母数の分布（E：指数分布, G：Gumbel分布, L：Logistic分布, N：正規分布）は母数に変化してもL-locationとL-scale

^{注9} ロジスティック分布および一般ロジスティック分布はわが国ではあまり用いられていない。

が変わるのみで3次以上のL積率は一定値であるので、図中では点で示されている。また、対数正規分布(LN3)とPearson type III分布(Pearson 3)は $\tau_3 = 0$ で正規分布(N)と重なる。LN3分布とGEV分布はGumbel分布のところで交差しておらず、Gumbel分布が対数正規分布のspecial caseであるという説は正しくない(角屋, 1956)。

従来、わが国ではL積率法(PWM法)の積率法に対するバイアスや誤差分散が小さいことが評価されてきているが、さらなる利点として、積率法ではAppendix 式(A3.9)に示すように標本歪み係数のとりうる限界値が標本サイズによって変化するので標本サイズに関係せずに種々の分布モデルを比較することができないのに対し、L積率法では図3.1のように種々の分布モデルのL-skewnessとL-kurtosisが一枚の図に表されるので、標本のL-skewness τ_3 とL-kurtosis τ_4 を知れば、概ねどの分布に近いかが判断できる利点がある。

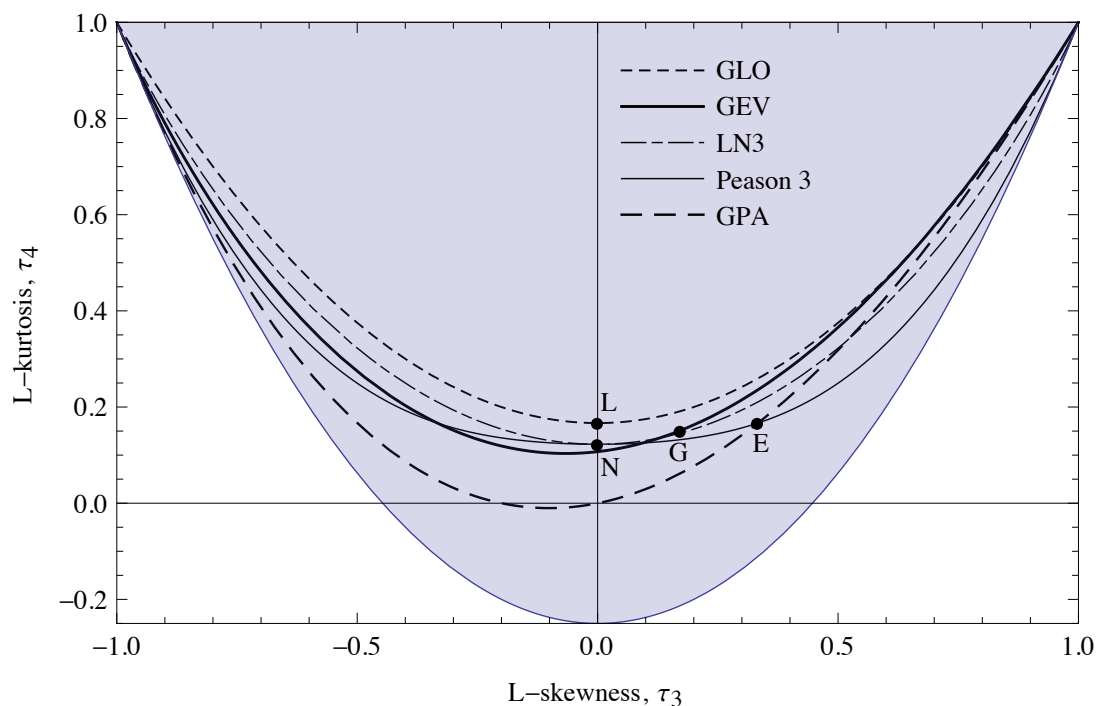


図3.1 各種分布のL-skewnessとL-kurtosisの比較

2母数の分布は点で、3母数の分布は線で示されている。着色部分は3次および4次のL積率の存在する範囲である。各記号は、それぞれ、E：指数分布、G：Gumbel分布、L：Logistic分布、N：正規分布、GLO：generalized logistic分布、GEV：一般極値分布、GPA：一般パレート分布、LN3：3母数対数正規分布、Pearson 3：Pearson type III 分布を示す

4) 平方根指数分布系のモデル：平方根指数分布，平方根指数型最大値分布

江藤ら(1986)により提案されている平方根指数分布および平方根指数型最大値分布の特性を検討するため、それぞれの分布モデルについて少し詳しく紹介する。

【平方根指数分布 (SQRT-ET分布)】

江藤・室田(1984)は一雨の総雨量の確率分布として、確率密度関数が次式で表される「平方根K分布」を提案している。

$$h(x) = \frac{2^{\nu-1} \beta}{\Gamma(1-\nu)} \left(\sqrt{\beta x} \right)^{-\nu} K_{\nu} \left(\sqrt{\beta x} \right) \quad (3.22)$$

ここに、 $K_\nu(s)$ は第2種変形ベッセル関数であり、次式で表される。

$$K_\nu(s) = \frac{1}{2} \left(\frac{s}{2} \right)^\nu \int_0^\infty t^{-\nu-1} \exp\left(-t - \frac{s^2}{4t}\right) dt \quad (3.23)$$

江藤ら(1986)は、式(3.98)の近似式として、

$$h(x) = \frac{\beta}{2} \exp\left(-\sqrt{\beta x}\right) \quad (x \geq 0) \quad (3.24)$$

を提案し、「平方根指数分布(SQRT-ET分布)」とよんでいる。この累積分布関数は次式で表される。

$$H(x) = 1 - \left(1 + \sqrt{\beta x}\right) \exp\left(-\sqrt{\beta x}\right) \quad (x \geq 0) \quad (3.25)$$

この分布には位置母数がなく、尺度母数 β のみの1母数の分布である。

x の分布範囲： $0 \leq x < \infty$

クオントイル：解析的な解はない。

$$\text{L積率：} \quad \lambda_1 = \frac{6}{\beta}, \quad \lambda_2 = \frac{15}{4\beta} \quad (3.26)$$

$$\tau_3 = \frac{121}{243} = 0.49794, \quad \tau_4 = \frac{4457}{15552} = 0.28659 \quad (3.27)$$

$$\text{母数推定法：} \beta = \frac{6}{\lambda_1}, \text{ または } \beta = \frac{15}{4\lambda_2} \quad (3.28)$$

平方根指数分布(SQRT-ET分布)は $0 \leq x < \infty$ の事象の発生確率を表すように考えられたものであり、位置母数がない。したがって、正の閾値以上の事象を対象にする場合には標本の各データから閾値を差し引いたものを対象に解析する必要がある。平均値である1次のL積率は閾値の値により変動するので、閾値が明白な場合はそれを用いればよいが、そうでない場合は閾値の大きさに影響されない2次のL積率から β を得る。次にこれを用いて $6/\beta$ とほぼ等しい1次L積率となる閾値を選べばよい。

図3.2は確率分布モデルのL-skewnessとL-kurtosisの関係図(図3.1)に式(3.27)で表される平方根指数分布のL-skewnessとL-kurtosisの値をクロス印で示したものである。平方根指数分布のL-skewnessが0.49794と指数分布の1/3に比べ大きく、右に大きく尾を引く分布であるのが特徴である。分布関数のL-skewnessは母数 β に依存しない。標本が有する歪みが理論値と同程度に大きい場合は適合度がよいが、歪みが小さいものや歪みが負の標本に対しては適合性はよくないと考えられる。

【平方根指数型最大値分布 (SQRT-ET max 分布)】

平方根指数分布で閾値以上の事象の発生確率が表され、その発生間隔がポアソン分布で表される場合、年最大値の確率分布は、式(3.29)のように表され、式(3.25)を代入して整理すると式(3.30)になる。江藤らはこの分布を平方根指数型最大値分布(SQRT-ET max 分布)と呼ぶこととしている^{注10}。

^{注10} わが国では年最大値を対象とした解析が行われることが多く、日本語では「平方根指数型最大値分布」と江藤らの表記のままよばれているが、英語表記では、水文・水資源ハンドブック(1997)、星(1998)、水理公式集(1999)などのように平方根指数分布の説明なしに平方根指数型最大値分布を単に“SQRT-ET”と呼んでいることが多いので注意する必要がある。

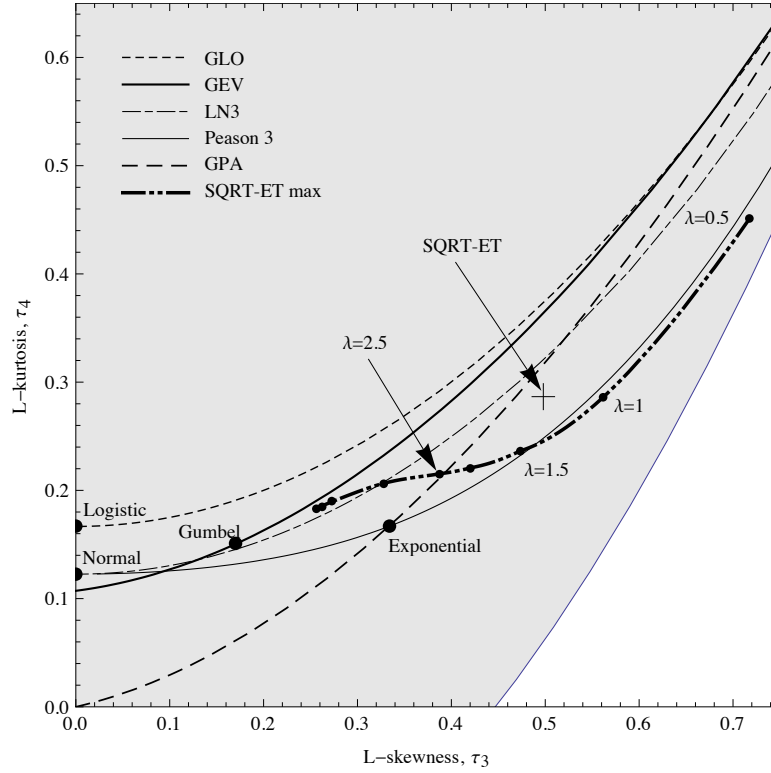


図3.2 平方根指数分布および平方根指数型最大値分布の τ_3 - τ_4 関係と他の分布との比較
 図中平方根指数型最大値分布上の点は右から順に $\lambda = 0.5, 1.0, 1.5, \dots, 80$ に対応している

$$F(x) = \exp[-\lambda \{1 - H(x)\}] \quad (3.29)$$

$$F(x) = \exp\left[-\lambda\left(1 + \sqrt{\beta x}\right) \exp\left(-\sqrt{\beta x}\right)\right] \quad (0 \leq x) \quad (3.30)$$

ここに、 λ は1年あたりの事象の発生回数である。なお、式(3.30)は $x=0$ において $e^{-\lambda}$ となり0にはならない。 $e^{-\lambda}$ は λ が7以上ではじめて0.001未満となる。些細なことかもしれないが、最大値の分布を議論する際に、 $x=0$ で $e^{-\lambda}$ となることは影響があると思われる。

確率密度関数は式(3.30)を x で微分して求められる。

$$f(x) = \lambda h(x) F(x) \quad (0 \leq x) \quad (3.31)$$

母数推定法として江藤ら(1986)は最尤法による解法を提案している。 λ は年発生回数であるが、毎年値資料では λ は不明であるので、尺度母数 β とともに最尤法により求めることとしており、データに最適化される。

なお、式(3.31)に対し、江藤ら(1986)は確率密度関数として

$$f(x) = \lambda h(x) F(x) / \{1 - \exp(-\lambda)\} \quad (0 \leq x) \quad (3.32)$$

と表記しているが、式(3.30)の x に関する微分は式(3.31)となる。なお、式(3.31)を区間 $[0, x]$ で積分すると

$$F(x) = \exp\left[-\lambda\left(1 + \sqrt{\beta x}\right) \exp\left(-\sqrt{\beta x}\right)\right] - \exp(-\lambda) \quad (0 \leq x) \quad (3.33)$$

となり、式(3.30)と $e^{-\lambda}$ だけ異なる。式(3.33)を x で微分すると式(3.31)が得られることで問題がなくなるが、式(3.30)は式(3.29)から導かれたものであり、導出過程に不明朗な部分が残る。

式(3.30)で表される分布のL積率を計算するのは容易ではないので、数値計算により求めた結果を示す。図3.3および3.4は実用的と思われる β および λ の範囲の1次L積率 λ_1 および2次L積率 λ_2 である。 λ の小さい領域ではそれ以上の領域に比べ関数形が大きく異なることがわかる。

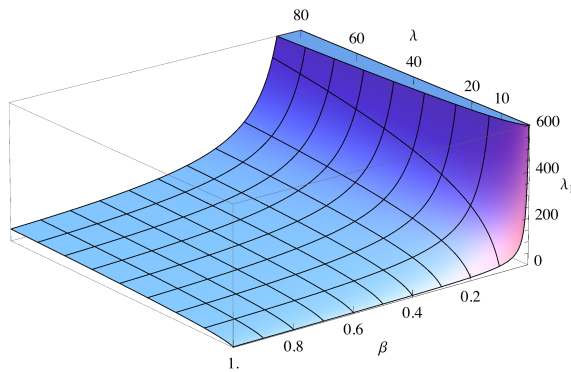


図3.3 平方根指数型最大値分布の1次L積率 λ_1

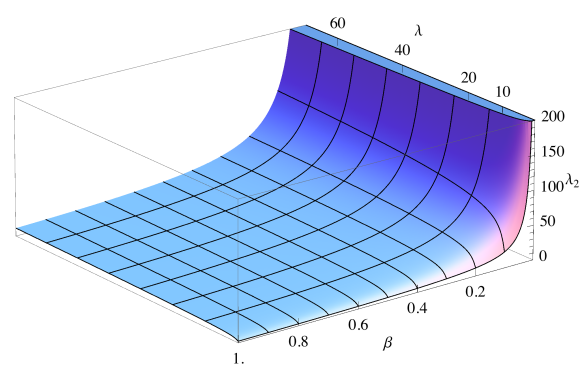


図3.4 平方根指数型最大値分布の2次L積率 λ_2

図3.5はL積率比 τ_3 および τ_4 と年発生回数 λ の関係を示す。L積率比 τ_3 および τ_4 は尺度母数 β の関数ではなく、年発生回数 λ のみの関数である。 λ が1付近で τ_3 および τ_4 は大きな値となり、 λ が大きい時、それらの比はほぼ0.7である。

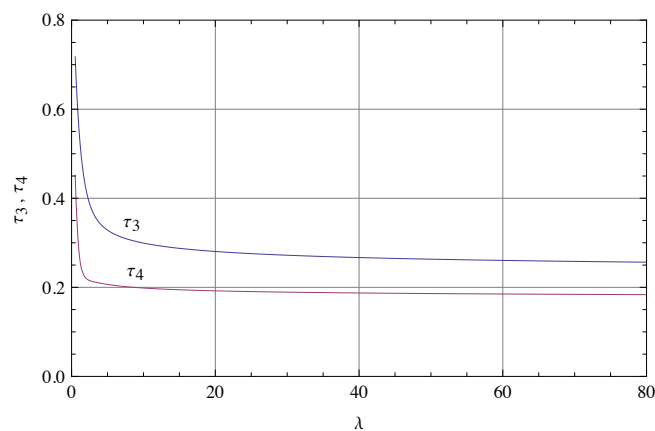


図3.5 τ_3 および τ_4 と年発生回数 λ の関係

前出の図3.2に $\lambda=0.5, 1.0, 1.5, \dots, 80$ について求めた $\tau_3-\tau_4$ の関係を他の分布と比較して示している。指数分布やGumbel分布等他の2母数の分布がこの図では点で表されるのに対し、平方根指数型最大値分布の場合 λ の値に応じて変化する。この図からわかるように平方根指数型最大値分布の $\tau_3-\tau_4$ 関係は、 λ が0.5~1.5の場合、Peason type III分布の値とほぼ同じであり、 $\lambda=2.5$ で一般パレート分布の値とほぼ同じになり、4以上になると対数正規分布と近くなり、さらに λ の値が大きくなるとGEV分布に近づく。江藤ら(1986)は λ が大きくなるとGumbel分布に漸近するかもしれないと述べているが、 λ が300でも τ_3 の値が0.242であり、GEV分布には近づいているもののGEV分布の特殊な場合であるGumbel分布に漸近するまでには至らない。本来、江藤らは λ が小さい場合は想定していないと思われるが、図3.5にみられるように、 λ が大きい場合に比べ λ の値の少しの変化に対し敏感に対応して変化

する。このため、図3.2で λ が0.5から2.5に変化する間に τ_3 は0.718から0.388まで大きく変化する。いずれにしても、平方根指数型最大値分布はGumbel分布より τ_3 が大きく、GEV分布における形状母数 k が負の領域にあり、上限をもつような標本($k>0$)やGumbel分布に適合するような標本($k=0$)にはうまく適合させることができない。

5) 平方根指数分布と平方根指数型最大値分布の関係

平方根指数型最大値分布が平方根指数分布から導かれているので、実際のデータを用いてこれらの関係をみてもみる。平方根指数分布は閾値以上の発生確率を表すものとして考えられるのでPOT(Peaks Over Threshold; 閾値以上の全データについて対象とするデータ系列)を用い、平方根指数型最大値分布には年最大値系列(AMS; Annual Maximum Series)を用いる。使用したデータは1980年から2004年までの25年間の四国、京上地点の日雨量のアメダスデータである。図3.6は、この標本から上位 $\lambda \times 25$ 個を $\lambda=1, 2, \dots, 143$ について取り出し、143セットのそれぞれの標本について式(3.28)の1次L積率を用いて求めた母数 β と2次L積率を用いた母数 β を比較したものである。後に示すようにこれらのデータの $\tau_3-\tau_4$ 関係が理論値と最も近い値となるのは $\lambda=35$ である。このときに、1次L積率を用いて求めた母数 β と2次L積率を用いた母数 β 、すなわち、平均値でみたものと分散でみたものがほぼ同じ母数 β を与えており、それぞれ $\beta=0.1874$ および $\beta=0.1885$ である。図中には、元の25年の年最大値系列を用いて、江藤らが提案している最尤法により求めた平方根指数型最大値分布の母数 $\beta=0.196$ をクロス点で示した。この事例ではかなり近い値となっている。しかしながら、平方根指数型最大値分布をAMS資料に適用して求められた λ は47となっており、35に比べかなり大きい。仮に平方根指数型最大値分布をAMS資料に適用して得られた $\lambda=47$ が適切とした場合、それに対応するPOT資料での母数 β の値は1次L積率を用いると $\beta=0.2134$ 、2次L積率を用いると $\beta=0.2116$ であり、少し異なる。また、実際上の問題として、分布関数が位置母数をもたないので、AMS資料から母数 β および λ が求まっても、これをPOT資料に変換して適用することができない。

$\lambda=35$ の場合、データの最小値は17mmとなるので、これを閾値として、これ以上のデータの発生を調べる。図3.7に示すようにヒストグラムと比べる限り平方根指数分布の適合性はよい。

図3.8は143セットのPOT標本について τ_3 と τ_4 を求めたものを図3.2に重ねてプロットしたものである。図中クロス点付近の○印が $\lambda=35$ に相当する。この $\lambda=35$ の標本の $\tau_3-\tau_4$ 関係が平方根指数分布の理論値に最も近くなっており、図3.7で示したように確率密度関数もよく度数分布に適合している。

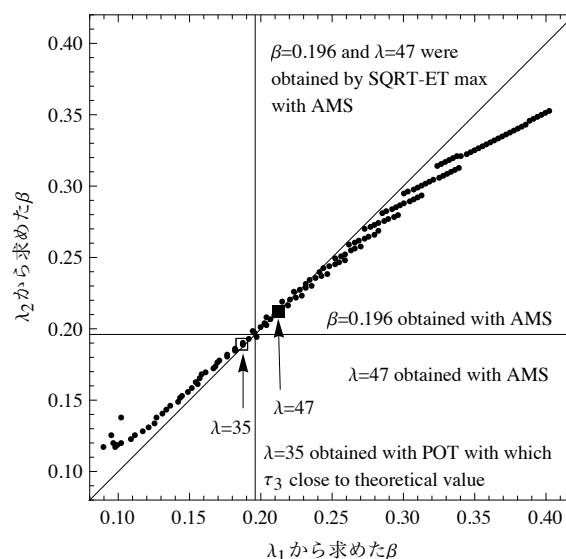


図3.6 1次L積率と2次L積率から求めた母数 β の比較

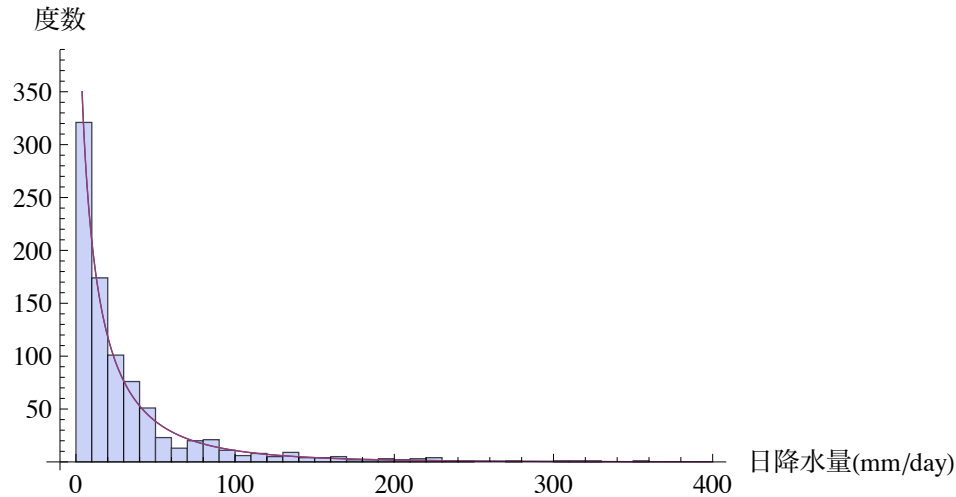


図3.7 平方根指数分布の日雨量データへのあてはめ
対象データは17mmを差し引いてプロットしている

一方、AMS資料について求めた τ_3 と τ_4 は図中“○”で示す値であり、 $\lambda=47$ の平方根指数型最大値分布の理論値より τ_3 および τ_4 が遥かに小さいデータであることがわかる。○印はGumbel分布より原点に近い側に位置しており、このAMS標本にGEV分布を当てはめると形状母数 k が正になり、上限を持つ分布となることが想定される。平方根指数型最大値分布の曲線上で $\lambda=47$ に対応する点を“□”で示している。 $\lambda=80$ とほとんど重なっている。○と□はかなり乖離しており、前述の形状母数で判断されるようにデータのプロットは上限をもつ傾向にあるにもかかわらず、平方根指数型最大値分布は常に右裾が長い分布となる。したがって、このAMS標本について平方根指数型最大値分布が適合度よく当てはまるとは考えられない。また、第4章でみるように流域平均雨量や洪水ピーク流量は右裾が長いデータ

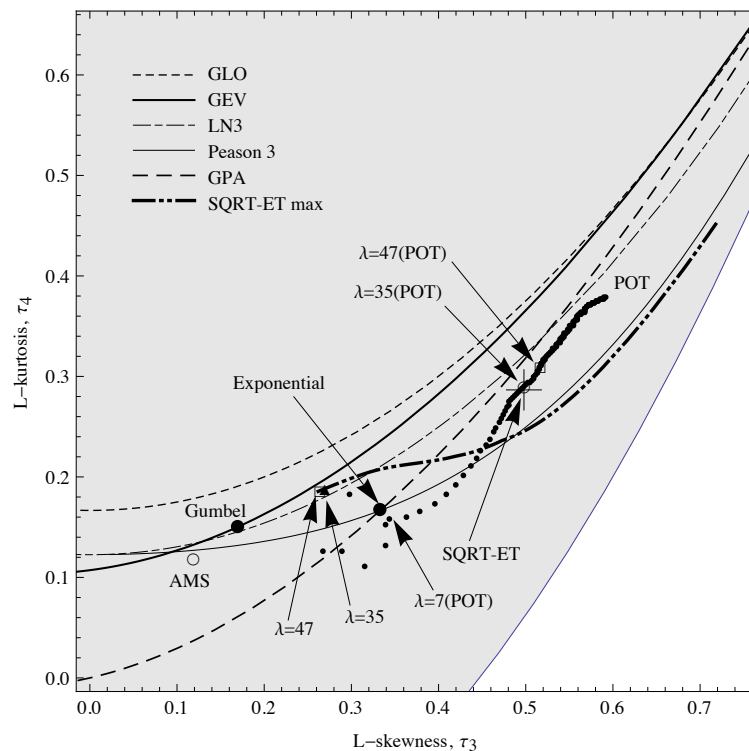


図3.8 四国京上地点のアメダス日雨量POT資料の τ_3 と τ_4 と理論値の比較

はそれほど多くないので、平方根指数型最大値分布が良好な適合度を示すものは多くない。なお、この図で興味深いのは、POT資料では λ が大きくなると τ_3 が大きくなるのに対し、AMS資料では逆に λ が大きくなると τ_3 が小さくなることであり、年間発生数 λ によりPOT資料とAMS資料の右裾が長い程度が異なることである。

AMS資料とPOT資料で仮に λ が互換的に適用できるとすると、図中POTのプロット上に”□”で示した値が $\lambda=47$ に対応するPOT資料の値である。逆に、POT資料の最適値である $\lambda=35$ に対応するAMS資料の値は▲である。

τ_3 と τ_4 の関係が指数分布のそれに最も近くなるPOT標本は $\lambda=7$ の場合であり、この場合の閾値は62mm/dayとなる。AMS標本の最小値が63mm/dayであるので、 $\lambda=7$ のPOT資料はAMS標本のすべてのデータを含む標本となっている。この状況をふまえるとこの場合、POT標本が指数分布、AMS標本がGumbel分布に良好に適合する可能性がうかがえる。

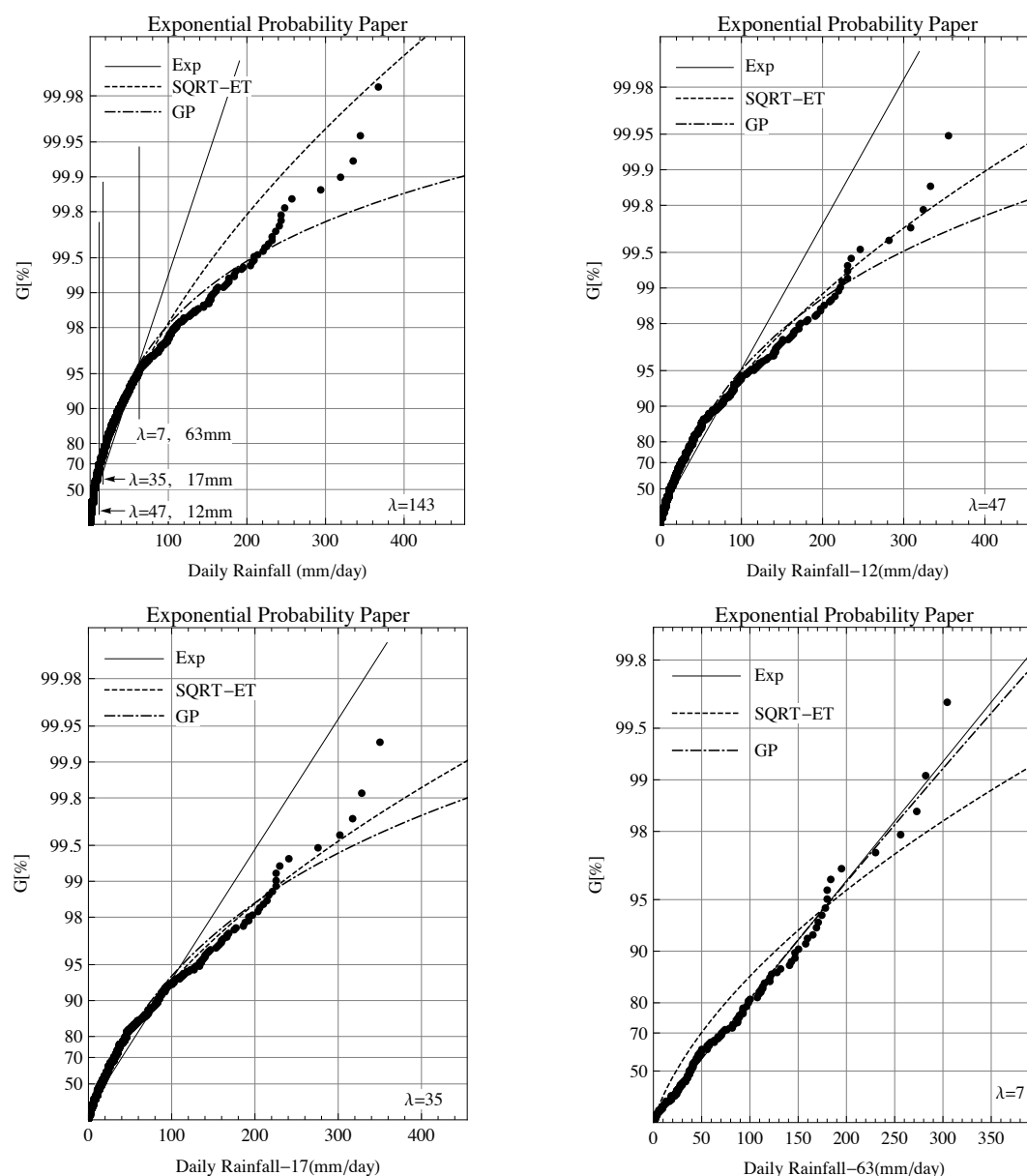


図3.9 京上地点アメダスデータのPOT標本への指数分布、一般パレート分布および平方根指数分布のあてはめ

図3.9に指数確率紙に $\lambda=143, 47, 35$ および7の場合の標本に指数分布，一般パレート分布および平方根指数分布をあてはめたものを示す。左上に示した $\lambda=143$ の図には他のグラフの閾値をあわせて示している。この図から， $\lambda=143, 47$ および35の場合，平方根指数分布が全体的によく適合しているのが分かる。指数分布は降雨量が小さい多くのデータに適合するように母数が決定され，超過確率が小さい部分での適合性は悪い。一方， $\lambda=7$ に対応する日雨量63mm以上のデータは指数確率紙上でほぼ直線的に分布しており，この部分のみを取り出した右下の図では指数分布および一般パレート分布の適合性が良好であり，平方根指数分布の適合性はよくない。

図3.10はAMS資料をGumbel確率紙にプロットするとともにそのデータにGumbel分布，GEV分布，平方根指数型最大値分布をあてはめたものであり，あわせてPOT資料も式(3.29)と同様に次式の変換をおこなってプロットしている。

$$p_{\text{AMS}} = \exp[-\lambda \{1 - p_{\text{POT}}\}] \quad (3.34)$$

ここに， p_{AMS} はAMS系列と比較できるように変換したプロットングポジション， p_{POT} はPOT資料系列におけるプロットングポジションである。

この図をみるとAMS資料およびPOT資料ともにGumbel分布またはGEV分布は良好にあてはめられているとみることができる。前述したようにAMS資料の τ_3 がGumbel分布の理論 τ_3 より小さいことから，GEV分布は上限値を持つような分布となつてあてはめられている。平方根指数型最大値分布はAMS資料のデータの並びに関係せず，上に凸のグラフとなっており，上限値を持たない分布となつていることがわかる。なお，この図の左下に注目すると，AMS資料の最小値と同じプロットングポジションとなるPOT資料のデータは大きい方から数えて94番目のデータである。したがって，これより

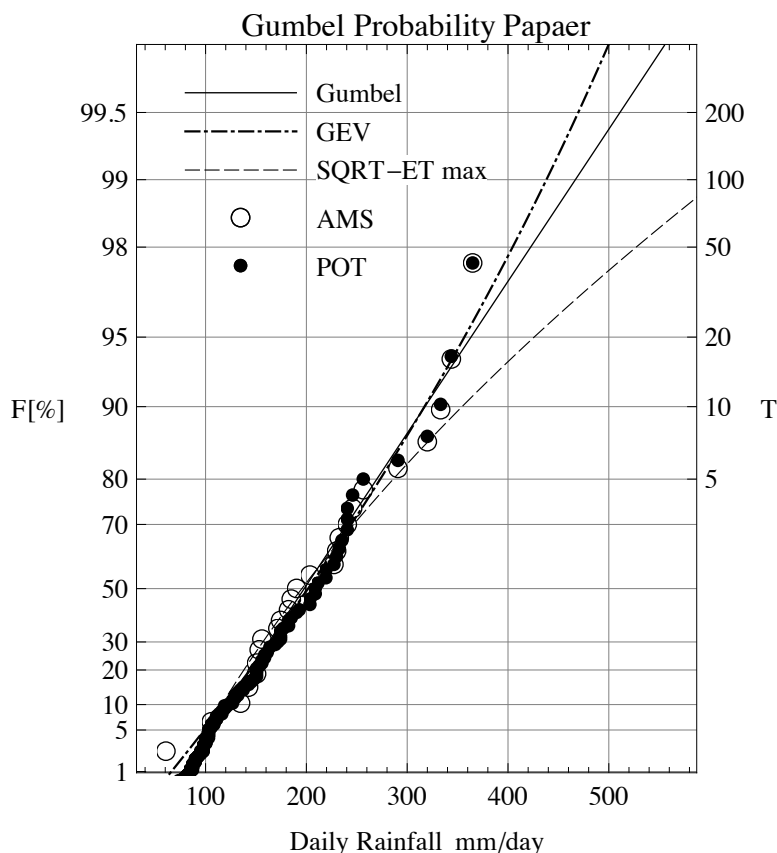


図3.10 京上地点アメダスデータのAMS資料への分布モデルのあてはめとPOT資料の比較

非超過確率の大きいデータはAMS資料が25個に対し、POT資料では約3.8倍あることになる。このデータ数の一般的な関係は式(3.34)にプロットングポジション式を代入して求めることができる。観測期間を N_y （AMS資料の標本サイズに等しい）、POT資料における標本サイズを n とすると、AMS資料の最小値に相当するプロットングポジションとなるPOT資料データの順位 i_1 は次式

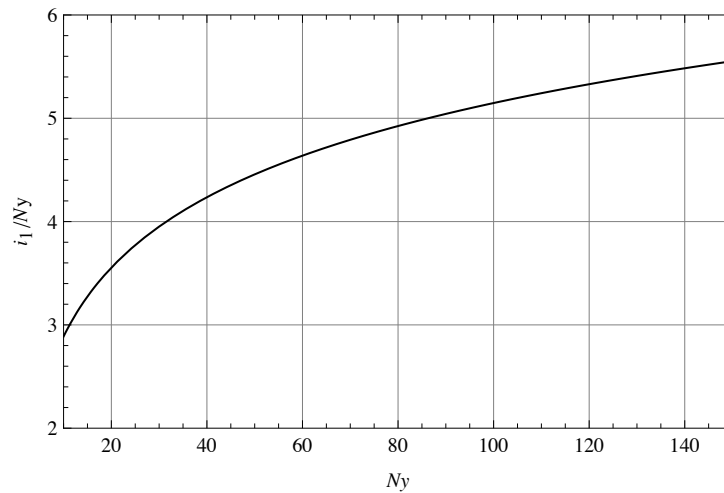


図3.11 AMS資料の最小値の非超過確率に相当するPOTデータの大きい方からの順位と N_y との比

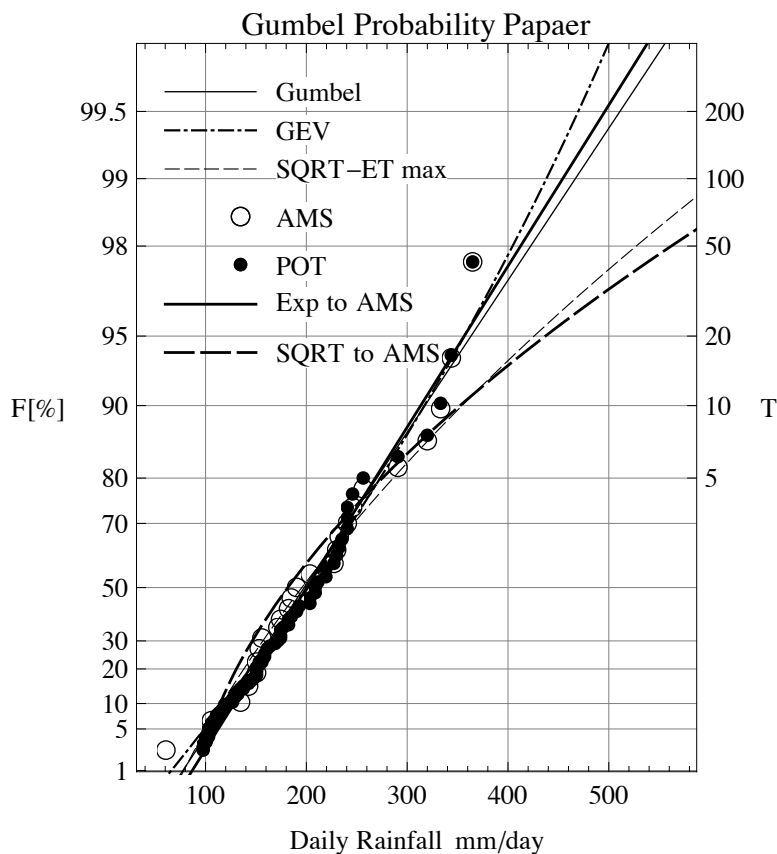


図3.12 京上地点アメダスデータのAMS資料への分布モデルのあてはめとPOTデータの比較

$$\frac{1-\alpha}{N_y+1-2\alpha} = \exp\left[-\frac{n}{N_y}\left(1-\frac{n+1-i_1-\alpha}{n+1-2\alpha}\right)\right] \quad (3.35)$$

を解いて、

$$i_1 = \alpha - \frac{N_y}{n} (1+n-2\alpha) \log\left[\frac{1-\alpha}{1+N_y-2\alpha}\right] \quad (3.36)$$

を得る。式(3.36)の関係を N_y との比で示したものが図3.11である。この関係は標本サイズが変化してもほとんど変化しない。図3.12はこのようにして得られた i_1 番目までのPOTデータとAMS資料系列をGumbel確率紙上で比較したものであり、あわせてこのPOTデータにあてはめた指数分布と平方根指数分布の母数を用いたGumbel分布および平方根指数型最大値分布を描いている。平方根指数型最大値分布はAMS資料より $\lambda=3.8$ のPOT資料の方が τ_3 の値が大きいので、AMS資料で母数推定したもの（図中SQRT-ET max）よりPOT資料を用いて母数推定したもの（図中SQRT to AMS）の方が結果的に図のように右に大きく尾を引く分布になっている。一方、指数分布については尺度母数がPOT資料を用いたもの（図中Exp to AMS）がAMS資料で母数推定したもの（図中Gumbel）より大きく（図では傾きが大きくなって）なっており、大きな事象への適合性が少しよくなっている。

このように、平方根指数や平方根指数型最大値分布はひずみが大きく裾の長い標本にはよくあてはまると考えられるが、そうでない場合にはデータと乖離した結果を与えるので注意が必要である。

3.3 推奨される母数推定法

寶(1989)は、3母数対数正規分布（母数推定法として、最尤法、積率法（標本歪）、積率法（不偏歪）、積率法（B-R歪）、石原・高瀬法、最小二乗法(Weibull)、最小二乗法(Hazen)、PWM法、岩井改良法、Sextile法）、Gumbel分布（最尤法、積率法、最小二乗法(Weibull)、最小二乗法(Hazen)、PWM法、最大エントロピー法）、GEV分布（最尤法、積率法（標本歪）、積率法（不偏歪）、最小二乗法(Weibull)、最小二乗法(Hazen)、PWM法）について、モンテカルロ法により発生・作成した5,000標本に対して、母数推定を行い、バイアスとRMSE(Root Mean Square Error)の観点から標本サイズ n 毎に推薦する母数推定法を検討している。その結果から、 $n=40$ あるいは50を念頭に整理すると以下ようになる。

対数正規分布： $30 \leq n \leq 100$ では、石原・高瀬法、積率法（B-R歪）がよさそうであるが、母数推定がうまくいかない場合がある。この点からは、PWM法がよい。

Gumbel分布： $30 \leq n \leq 500$ ではPWM法がよい。

GEV分布： $n \leq 500$ ではPWM法がよい。

3.4 候補モデルのscreening方法

治水計画では、一般にたまにしか起こらない規模の大きい現象を対象とし、年最大値などの極値資料から100～200年に1回起こるといような非常に低い確率の現象を推定することが求められる。極値に種々確率分布をあてはめたときの最大の問題は、極値分布の裾付近で観測資料が適合しているかどうかである。この分布裾での適合度が計画の信頼性を決定してしまうからである。確率水量量是用いた標本の大きさ、分布形及び母数推定法により大きく変動することはよく知られており、これが常に計画策定者を悩ませる問題でもある。（星，1998）

候補モデルのscreeningは適合度の悪いモデルを候補から除外するものである。適合度を測る規準としては相関係数、標準最小自乗規準(SLSC)、最大対数尤度、赤池の情報量規準(AIC)等がある。

同じ長さの変量ベクトル X と変量ベクトル Y があるとき、相関係数 ρ_{xy} は

$$\rho_{xy} = \frac{(X - \mu_X) \cdot (Y - \mu_Y)}{\sqrt{\{(X - \mu_X) \cdot (X - \mu_X)\} \{(Y - \mu_Y) \cdot (Y - \mu_Y)\}}} \quad (3.37)$$

で表される。ここに、 μ_X および μ_Y はそれぞれ X および Y の平均であり、 $X - \mu_X$ はベクトル X の各要素から平均値 μ_X をそれぞれ引いたベクトルを意味する。また、 \cdot は内積を表す。相関係数を考える場合、確率でみるものと変量で見るものの2つが考えられる。前者は順序統計量のそれぞれの値を確率分布モデルで評価した非超過確率と順序統計量のプロットングポジションから求まる非超過確率との関係を見るものであり、後者はプロットングポジションから求まる非超過確率に対応するように確率分布モデルから計算されるクォンタイルと順序統計量がどのような相関を有するかをみるものである。以下では前者をPCOR、後者をXCORと呼ぶこととする。相関係数は1に近いほど適合度がよいと判断される。

宝・高棹(1988)は適合度の判断規準として標準最小自乗規準(SLSC ; Standardized Least Square Criterion)を提案し、相関係数と比較した上でSLSCの利用を推奨している。

$$SLSC = \frac{\sqrt{\xi^2}}{|s_{99} - s_{01}|}, \quad \xi^2 = \frac{1}{n} \sum_{i=1}^n (s_i - s_i^*)^2 \quad (3.38)$$

ここに、 s_{99} および s_{01} はそれぞれ非超過確率0.99と0.01に対する当該確率分布の標準変量、 n ：標本の大きさ、 s ：順序統計量を推定母数で変換した標準変量ベクトル、 s^* ：確率分布モデルで求めたプロットングポジションに対応するクォンタイルを推定母数によって変換した標準変量ベクトルである。SLSCは値が小さいほど適合度が良いと判断される。

最大対数尤度は次式で表される。

$$MLL = \sum_{i=1}^n \text{Log}[f(X_i, \hat{\theta})] \quad (3.39)$$

ここに、 $f(X_i, \hat{\theta})$ は最尤法により求められた母数ベクトル $\hat{\theta}$ を用いて計算される確率変量 X_i の確率密度である。なお、Logは自然対数である。最大対数尤度は値の大きい方が適合度がよいと判断される。

赤池の情報量規準(AIC)は次式で表される。

$$AIC = 2m - 2MLL \quad (3.40)$$

ここに m は確率分布関数の母数ベクトルの長さ(母数の数)である。一般に母数の数が増えると分布の適合度はよくなるが、AICは母数の数を考慮していることが他の規準とは異なる特徴である。AICの値が小さいほど適合度がよいと判断される。

従来、河川計画においては、確率水文量の推定にあたって用いる確率分布の選択は、最小二乗法や上位10点最小二乗法、確率紙へのプロットの視覚的判断などが用いられてきた。しかしながら、これらは、主観的判断によるものであったり、確率分布をどれかに決定した後に最適化を行うものであるなど、適合度の客観的判断に疑問が残る。宝・高棹(1988)は適合度の絶対的な判定規準として標準最小二乗規準(SLSC)を提案している。適合度の判定規準には、他に相関係数や最大対数尤度、赤池の情報量規準などがある。最大対数尤度と赤池の情報量規準は、比較対象の確率分布の中での相対的評価

しかできない。相関係数よりSLSCの方が規準の持つ意味が明確である。このようなことを踏まえ、本研究では、SLSCにより適合度の評価を行うこととする。

3.5 確率水文量の変動性の評価方法

一般に、確率分布モデルは母数の個数が多くなるにつれて標本によく適合するようになるため、適合度の観点からは母数の個数が多い確率分布の方が適合していると判断されることが多い。しかしながら、このような確率分布は、新しくデータが加わった際に、非超過確率の大きい部分の確率水文量の変動が大きくなる可能性がある。流域の長期的な管理の目標の基礎となるものがデータの追加によってあまり変動するのも困るし、そもそもその確率水文量に対する信頼が揺らいでしまう。このため、単に適合度がよい確率分布ではなく、データの追加に対しても安定的な確率分布を選択する必要がある。星(1998)も水文資料の蓄積が進んでいる現在、標本数の違いによって確率水文量が大きく変動しないような確率分布モデルと母数推定法を選定することが望ましいと述べている。

本研究ではリサンプリング手法を用いて安定性の評価を行うこととする。

リサンプリング手法にはjackknife法やbootstrap法がある(Efron, 1982)。jackknife法は大きさ n 個の標本のうち i 番目の1データのみを欠いたデータ数 $n-1$ 個の標本をすべての i について作成し(n セット作成することになる)、これらの標本から求めた統計量をもとに不偏推定値およびそのまわりの推定誤差を算定する手法である。一方、bootstrap法は大きさ n 個の標本から重複を許して任意に n 個取り出した標本を複数作成し、これらの標本から求めた統計量をもとに不偏推定値およびそのまわりの推定誤差を算定する手法である。jackknife法は計算回数が少なく、作成する標本数および不偏推定値、推定誤差が一意的に定まるのに対し、bootstrap法は作成する標本数が任意に設定でき、作成する標本数によって不偏推定値や推定誤差が異なる。両手法によるバイアス(偏倚)補正量は、bootstrap法のバイアス補正量がjackknife法のバイアス補正量の $(n-1)/n$ となる(Efron, 1982)が、 n が極端に小さくなければこの差は小さい。このため、今回の検討における確率分布モデル相互の安定性評価の指標としてはjackknife法を用いる。

1) jackknife法の概要

Quenouille(1949)は統計量のバイアスを少なくする方法として標本を二分した標本に基づく方法を提案し、後に、サイズ h の g グループに分割する方法に一般化した($n=gh$)。その後、グループの選択可能性を排除した $h=1$ すなわち $g=n$ の特殊な場合についての研究が数多く行われ、これが普及している(Miller, 1974)。以下、 $h=1$ の場合について述べる。なお、“jackknife”とは“幅広いいろいろな状況の中で有用な大雑把な道具”という意味合いで付けられたものである。(寶, 1989)

jackknife法は大きさ n の標本の各データを x_1, x_2, \dots, x_n とする。これを用いて求める母集団の特性を推定する統計量を

$$\hat{\psi} = \psi(x_1, x_2, \dots, x_n) \quad (3.41)$$

とする。

大きさ n 個の標本のうち i 番目の1データのみを欠いたデータ数 $n-1$ 個の標本を用いた統計量を

$$\hat{\psi}_{(i)} = \psi(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \quad (3.42)$$

とする。 $\hat{\psi}_{(i)}$ は $i=1, 2, \dots, n$ に対して求まるので n 個求まることになる。

$\hat{\psi}_{(i)}$ の平均値を

$$\hat{\psi}_{(\cdot)} = \frac{1}{n} \sum_{i=1}^n \hat{\psi}_{(i)} \quad (3.43)$$

により求める.

Quenouilleのバイアス推定値は次式で与えられる.

$$\text{BIAS} = (n-1)(\hat{\psi}_{(\cdot)} - \hat{\psi}) \quad (3.44)$$

式(3.44)を用いて式(3.41)の統計量のバイアスを補正した”jackknife”推定値 (以後, jackknife推定値という) は次式で与えられる.

$$\tilde{\psi} = \hat{\psi} - \text{BIAS} = n\hat{\psi} - (n-1)\hat{\psi}_{(\cdot)} \quad (3.45)$$

また, jackknife法による推定誤差分散は,

$$\text{VAR} = \frac{n-1}{n} \sum_{i=1}^n (\hat{\psi}_{(i)} - \hat{\psi}_{(\cdot)})^2 \quad (3.46)$$

で求められる.

式(3.41)の統計量が $1/n$ のオーダーのバイアスを有しているのに対し, jackknife推定値のバイアスは $1/n^2$ のオーダーであり(Efron, 1982), jackknife法によるバイアスの補正が有効である.

本研究では確率水文量の安定性を評価するため, ψ に確率水文量を取り, 式(3.46)の平方根で表される推定誤差を安定性の評価指標として用いる.

2) Bootstrap法の概要

實(1989)によるとbootstrap法の概要は以下のとおりである. まず, 標本サイズ n のデータ x_1, x_2, \dots, x_n から繰り返しを許して n 個取り出し, それを $x_1^*, x_2^*, \dots, x_n^*$ と記す. この 1 組の標本を”bootstrap標本”という. bootstrap標本を用いて統計量を求め, それを次のように記す.

$$\psi^* = \psi(x_1^*, x_2^*, \dots, x_n^*) \quad (3.47)$$

以上の作業を多数回 (B 回) 繰り返す. すなわち, 全部で B 個のbootstrap標本それぞれに対して ψ^* を求める. 第 b 番目のbootstrap標本に対して得られた統計量を便宜上 ψ^{*b} と記す ($b=1, 2, \dots, B$).

次式により ψ^{*b} の平均値 $\psi^{*\cdot}$ を求める.

$$\psi^{*\cdot} = \frac{1}{B} \sum_{b=1}^B \psi^{*b} \quad (3.48)$$

これは統計量 ψ のbootstrap推定値である.

統計量 ψ の分散のbootstrap推定値は次式により求められる.

$$\text{VAR}_B = \frac{1}{B-1} \sum_{b=1}^B (\psi^{*b} - \psi^{*\cdot})^2 \quad (3.49)$$

3.6 Mathematicaを用いた水文統計解析プログラム

Mathematicaは1988年にWolfram Researchが開発・リリースした技術・科学計算のための統合ソフトウェア環境で、数値計算、数式処理、視覚化、プログラミング機能を1つの柔軟でインタラクティブなソフトウェアシステムに統合している。現在はversion 7となっており、積分等の計算や図形描画等の機能が大幅に改善されており、他の人が開発したプログラムを実行することができる簡易な無料版もある。

筆者が水文統計の研究をはじめた1997年当時、河川計画に関係する便利な水文統計ソフトウェアが見当たらなかったので開発する必要性を痛感した。筆者はFortranの知識はあったが、挙動等が明確でない関数を対象に解を求めるソフトウェアをプログラミングするのは不安であった。このような場合に試行錯誤を繰り返しているといつ完成するのか分からなくなってしまう。短期間に必要なソフトウェアを開発する必要性を感じていたなかで、MathematicaはかつてのBasic言語のようなインタープリタ的な性格を持ちながら、数式計算、グラフ描画等かなり先進的なソフトウェアであったので、これを用いれば短期間に開発することも不可能ではないと判断した。Mathematicaは一連のまとまった手続きをモジュールとして定義でき、それをあたかも関数を使うごとく実行することができる。これはこれを含む一連の手続きの中でも使えるし、単独の命令としても使える。モジュールを分かりやすく定義することが使いやすいプログラムになる。また、統計関係の計算では数表を用いることが多いが、プログラムに数表を参照させるのはあまりスマートな方法ではないと考えていたので、正規分布のクォンタイルが簡単に計算できたり、積分範囲を ∞ まで取ることが容易にできることは短期間の開発のめどを与えてくれるものであった。なお、現在では統計計算用には“R”という無料ソフトウェアが有名になっているが、これが開発されたのは1990年代半ばのことで1997年頃から利用できるよう環境が整っていたようであるが当時は分からなかった。

プログラムは以下のモジュールからなる。扱う分布としては、対数正規分布（岩井改良法、石原・高瀬法、クォンタイル法）、Gumbel分布、一般極値分布、平方根指数型最大値分布、対数ピアソンⅢ型分布、指数分布、一般パレート分布などについて作成した。

- ・時系列の洪水イベント資料からAMS資料およびPOT資料を取り出すモジュール
- ・L積率を求めるモジュール
- ・母数を求めるモジュール（形状母数を与えて残りの母数を求めるものも含む）
- ・リターンピリオドに対応した確率水文量を求めるモジュール
- ・変量に対応するリターンピリオドを求めるモジュール
- ・適合度をみるモジュール
- ・確率紙（対数正規確率紙、Gumbel確率紙、指数確率紙）にプロットするモジュール
- ・リサンプリング(jackknife,bootstrap)関係のモジュール
- ・POT資料をAMS資料と比較できるように変換してプロットするモジュール
- ・Q-Qプロットを作成するモジュール
- ・ヒストグラムを作成するモジュール

なお、本論文に使用している統計関係の図はすべてこのMathematicaによるプログラムおよびMathematicaの機能を用いて作成している。主に参考とした文献は水文統計論（角屋，1964），応用水文統計学（岩井・石黒，1970），Handbook of Hydrology(Stedinger et al., 1993)，大雨の頻度（江藤ら，1986），水文頻度解析における確率分布モデルの評価規準（宝・高棹，1988），水文頻度解析（星，1998）である。

3.7 結論

本章では、水文頻度解析の一般的手法について述べるとともに、江藤ら(1986)が提案した平方根指数分布および平方根指数型最大値分布の特性について調べた。この結果、平方根指数型最大値分布の導出過程において不明瞭な部分があることおよび累積分布関数を独立変数について微分しても提案されている確率密度関数にはならないことが分かった。このことは、極値分布には3つの型しか存在し得ないことおよびそれは一般極値分布で統一的に表されるという従来の研究に矛盾せず、平方根指数型最大値分布を極値分布として扱うことができないことを示している。

また、平方根指数分布および平方根指数型最大値分布はひずみが大きい資料には適合するがひずみが小さい資料には適合しないことが分かった。さらに、POT資料と平方根指数分布は理論的にもよくあうものがあるが、AMS資料およびAMS資料に対応するPOT資料については理論値と実資料はかなり乖離することが分かった。これらの分布を適用するにあたっては資料の歪み特性等を十分に調べた上で適用の可否を検討すべきである。単に適合度基準の数値のみで判断することは危険であり、確率紙にプロットしてデータの特性を十分把握したうえで適用すべきである。なお、L積率比 $\tau_3-\tau_4$ 関係で見ると、この分布で再現される範囲の歪みに対しては一般パレート分布やピアソンⅢ型分布、一般極値分布などがカバーできる範囲にある。

3.6節では本研究を遂行する際に開発・利用してきたMathematicaを用いた水文統計処理プログラムの概要を述べた。

参考文献

- Efron, B.: The Jackknife, the bootstrap and Other Resampling Plans, SIAM Monograph, No.38, p.92, 1982.
- Greenwood, J.A., J.M. Landwehr, N.C. Matalas and J.R. Wallis: Probability Weighted Moments —Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form, Water Resources Research, Vol. 15, No.5, pp.1049-1054, 1979.
- Hosking, J.R.M. and J.R. Wallis : Regional Frequency Analysis –An Approach Based on L-Moments–, Cambridge Univ. Press., p.207, 1997.
- Miller, R.G.: The jackknife –a review, Biometrika, Vol.61, No.1, pp.1-15, 1974.
- Quenouille, M.H.: Approximate tests of correlation in time-series, J.R. Statist. Soc. B 11, pp. 68-84, 1949.
- Stedinger, J.R., R.M. Vogel, and E. Foufoula-Georgiou: Frequency Analysis of Extreme Events, Chap. 18, Handbook of Hydrology, (Ed.) D. R. Maidment, McGraw-Hill, New York, pp. 18.1-18.66, 1993.
- 岩井重久, 石黒政儀：応用水文統計学, 森北出版, p.370, 1970.
- 江藤剛治, 室田 明：一雨降雨の1確率模型, 土木学会論文集, 第345号/II-1, a) p.103, b) pp. 104-105, 1984.
- 江藤剛治, 室田 明, 米谷恒春, 木下武雄：大雨の頻度, 土木学会論文集, 第369号/II-5, pp. 165-174, 1986.
- 角屋 睦：極値分布とその一解法, 農業土木研究, 第23巻, 第6号, pp.350-357, 1956.
- 角屋 睦：水文統計論, 水工学シリーズ, 土木学会水理委員会, 64-02, p.59, 1964.
- 寶 馨：水資源システムにおける確率論的モデルと手法の評価に関する研究, 京都大学博士論文, 1989.
- 宝 馨, 高棹琢馬：水文頻度解析における確率分布モデルの評価規準, 土木学会論文集, 第393号/II-9, pp.151-160, 1988.
- 星 清：水文頻度解析, 水文・水資源ハンドブック, 7.3, 水文・水資源学会編, 朝倉書店, pp. 238-248, 1997.
- 星 清：水文統計解析, 開発土木研究所月報, No.540, pp.31-63, 1998.

4. 年最大値系列を用いた水文頻度解析

本章では、年最大値系列(AMS)資料を用いた水文頻度解析結果について述べる。最初に、わが国の一級河川の流域平均雨量について解析する。次に一級河川の洪水ピーク流量について、適合度および安定性について評価し、モデルのscreeningの際の適合度判定基準についてのべる。

4.1 豪雨極値データの解析

1997年に河川法が改正され、まず、全国の一級河川において河川整備計画を策定すべく、河川整備基本方針の検討が開始された。全国の一級河川のうち、多くの河川で基準点上流の流域平均雨量が整理された。この作業において、流域面積の大きい河川では比較的長い短時間雨量について、流域面積の小さい河川では比較的短い短時間雨量について整理された。これらを概観すると、3時間、6時間、12時間、24時間での整理がかなり多くの河川で整理されているので、この4つの短時間雨量について頻度解析をおこなう。

前章で示したようにPOT資料はその発生がポアソン分布にしたがうときPOT資料の累積分布 $H(x)$ と年最大値の累積分布 $F(x)$ は次式の関係があり、AMS資料と比較できる。

$$F(x) = \exp[-\lambda \{1 - H(x)\}] \quad (4.1)$$

$H(x)$ が指数分布、一般パレート分布および平方根指数分布の場合、 $F(x)$ はそれぞれGumbel分布、GEV分布および平方根指数型最大値分布で表される。POT資料について考えるとき、指数分布は一般パレート分布の形状母数 k が0の場合に相当することおよび平方根指数分布に比べて一般パレート分布の方が図3.8で示したように掘り広い範囲に適用でき、柔軟である。したがって、その年最大値を考えるにあたっては一般パレート分布に対応するGEV分布を用いて全国のAMS資料を対象にした検討を行う。なお、対象とした資料は1999年時点で建設省の地方建設局（現在の国土交通省地方整備局）から資料整理の途中段階で提供いただいたものであり、必ずしもすべての一級河川の資料が整っているわけではないが全国的な傾向の把握はできると考えられる。

Gumbel分布はGEV分布の形状母数がゼロの特別なケースであり、指数分布は一般パレート分布の形状母数がゼロの特別なケースである。ここではAMS資料を対象とするので、GEV分布をあてはめ、その形状母数の特性についてみる。

各地方の河川について、水系名河川名、観測地点名、観測地点上流の流域面積、AMS資料の蓄積状況（24時間降雨量の観測年数）と3時間、6時間、12時間、24時間の流域平均雨量のGEV分布の形状母数 k をAppendix A-5に示した。一部の河川では3時間、6時間および12時間の観測年数が24時間のものと異なるものもあるが概ね時間雨量が観測され始めてからはこのような整理が可能となっている。また、図4.1はその形状母数 k を各地方ごとにまとめてプロットしたものである。図中、太線は各地方における各時間毎の k の値の平均値である。平均値で見ると、北から南へ向かうにしたがって徐々に k が負から正に転じて行く様子がわかる。すなわち、これはGEV分布を当てはめた場合、北では右に長く尾を引く分布になるのに対し、南へ向かうにしたがいGumbel分布に適合するようになり（北陸地方）、さらに南では上限を持つような分布に適合するようになることを示している。

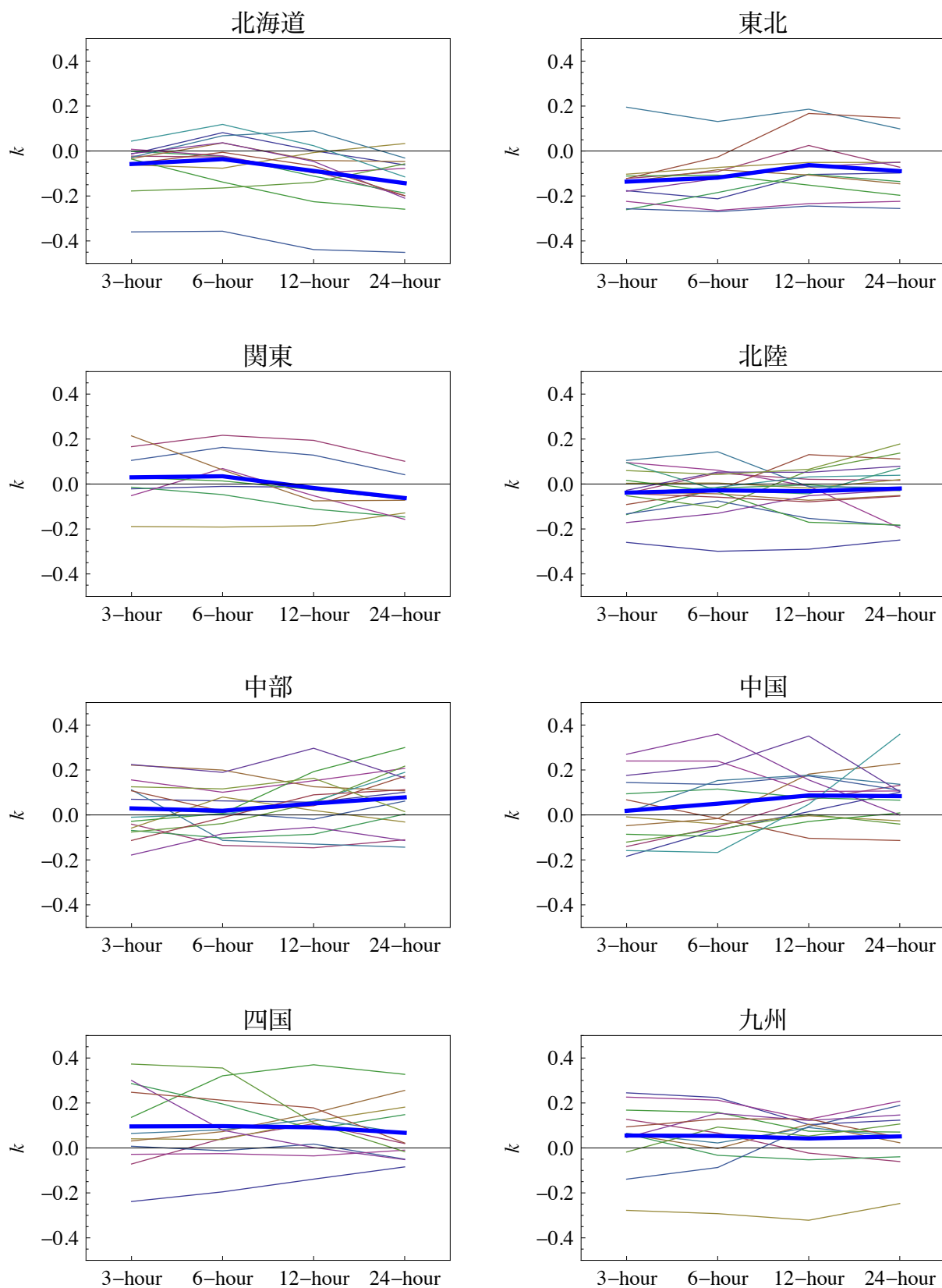


図4.1 全国のAMS資料のGEV分布の形状母数 k

各地方における平均を太線で示している。形状母数 k の値がが北から南へ向かうにしたがって徐々に負から正へと変化している

図4.1に示した各地方における k の平均値を3時間、6時間、12時間、24時間の雨量について示したものが図4.2である。いずれの時間雨量についても北の地方から南の地方に移るにつれて k の地域内平均値が四国地方まで徐々に大きくなる傾向があり、特に12時間雨量ではこの傾向が明確に現れている。また、24時間雨量では中部および四国をピークに四国および九州へ向かって少し小さくなる傾向にある。図4.3にはこれら4つの時間区分の平均値を示す。隣接する地域で k の平均値が少し逆転する場合もあるものの概ね北から南へ向かって k の平均値が負から正へ徐々に転換している。

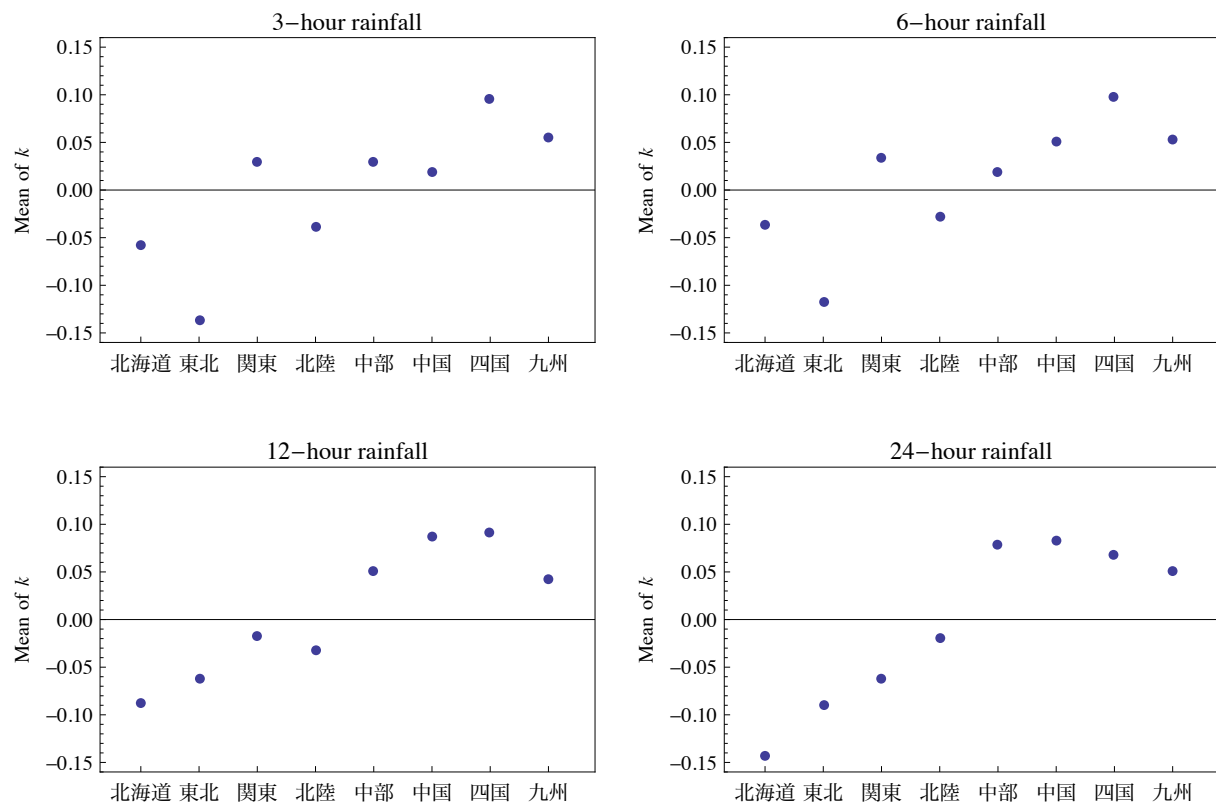


図4.2 各地方における3時間、6時間、12時間、24時間の流域平均雨量の k の平均値

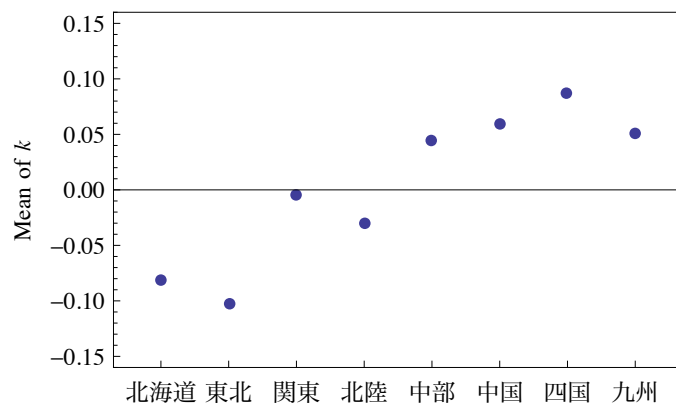


図4.3 各地方における流域平均雨量の k の平均値（河川および時間区分）

図4.4に東北地方において3時間、6時間、12時間、24時間の流域平均雨量の k の時間および空間にわたる平均値 -0.102 と一地点における3時間、6時間、12時間、24時間の k の平均値が最も近い高瀬川の3時間、6時間、12時間、24時間の流域平均雨量のAMS資料をGumbel確率紙にプロットしたものを示す。図4.5および図4.6は同様にそれぞれ北陸地方の関川水系保倉川および四国地方の吉野川について示したものである。なお、四国地方については瀬戸内海側の河川は太平洋側の河川と異なる特徴を有していると考えられるのでこれらを除いて考えることとし、太平洋側の河川の k の平均値は 0.121 となるのでこれに最も近い吉野川を選んだ。これを見ると形状母数で右に長く尾を引くか上限を持つのか判断できるのでGumbel確率紙上でのプロットの傾向が大方判断ができることが理解できる。

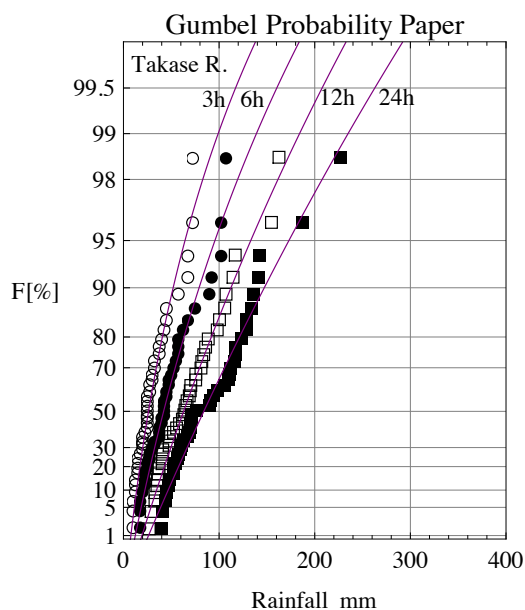


図4.4 東北、高瀬川の3h,6h,12h,24hのAMS資料とGEV分布のあてはめ

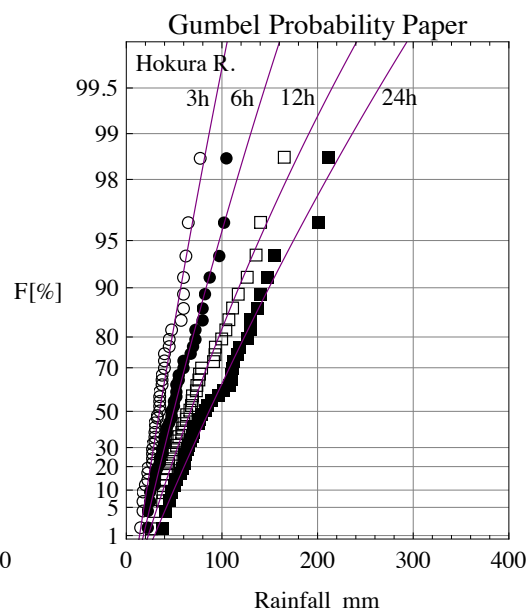


図4.5 北陸、保倉川の3h,6h,12h,24hのAMS資料とGEV分布のあてはめ

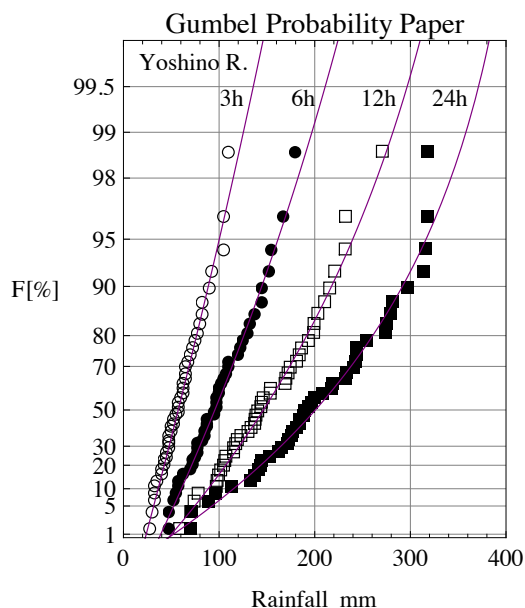


図4.6 四国、吉野川3h,6h,12h,24hのAMS資料とGEV分布のあてはめ

なお、一般的に治水計画に用いる降雨量の降雨継続時間はいろいろなパターンの実績降雨が取り込まれるよう24時間またはもっと長い時間が採用されることが多い。しかしながら、洪水到達時間は24時間以内のものが多く、洪水到達時間内の豪雨がその河川の洪水ピークと密接に関係すると考えられることから、降雨継続時間よりも洪水到達時間に注目した解析検討も必要である。とくに、GEV分布の形状母数が正になり、一見豪雨の降雨量に上限があるような解析結果となるところでは、検討の価値はある。角屋(1964)は対数極値分布（GEV分布に相当する）適用の際の実用上の問題について、対数極値分布B型（GEV分布で形状母数が正になり、上限がある）の場合、危険側の答えを出すことになって不都合であるのでGumbel分布を適用することにするとして述べている。全国の中部から南西部にかけての地方で図4.6のようなパターンになる河川では計画検討にあたり単に解析結果の数値のみにとらわれずに降雨継続時間より短い洪水到達時間に注目した解析検討も行なって判断する必要がある。

図4.7は流域面積とGEV分布の形状母数の関係を地域別に示したものである。流域面積が小さい部分では流域によりいろいろな k の値をとっているが、流域面積が大きい部分ではこのばらつきが小さくなっている。関東および中国を除き、わずかではあるが、いずれの地域でも流域面積が大きくなるとGEV分布の形状母数が大きくなる傾向にある。すなわち、ある地域の平均的な形状母数に対し、流域面積が大きくなれば少し正の側にシフトする。これは大きな流域ほど既往のデータが上限を持ちやすい傾向にあることを示している。なお、ここで扱ったデータの範囲ではGumbel分布に近づくという見方もできる程度である。

しかしながら、わが国の最大降水量の記録は世界の記録に比べると小さい（水文・水資源ハンドブック、1997）。世界の記録と日本の記録は条件が異なり、必ずしもそのまま比較できるとは考えられないが、これまでの年最大値がいかにも上限値を有するようなプロットとなっても、これは必ずしも物理的な上限の存在を示しているわけではなく、いずれ未曾有の大雨に遭遇することもありうると考えていろいろな観点から吟味すべきである。たとえば、図4.6に示した吉野川の12時間のプロットを見ると、2位以下のデータの並びと比べ、かなり大きな値となっている。このため、上述の角屋(1964)の示した方向性は示唆に富むものである。なお、この場合、適合度のよいものよりも適合度はよくないが安全側のものを選ぶということになり、頻度解析の結果を重視する立場からは恣意的とみられることもあるので、水文頻度解析の結果を計画論にどのように活かすかという考え方を明確に整理しておく必要がある。

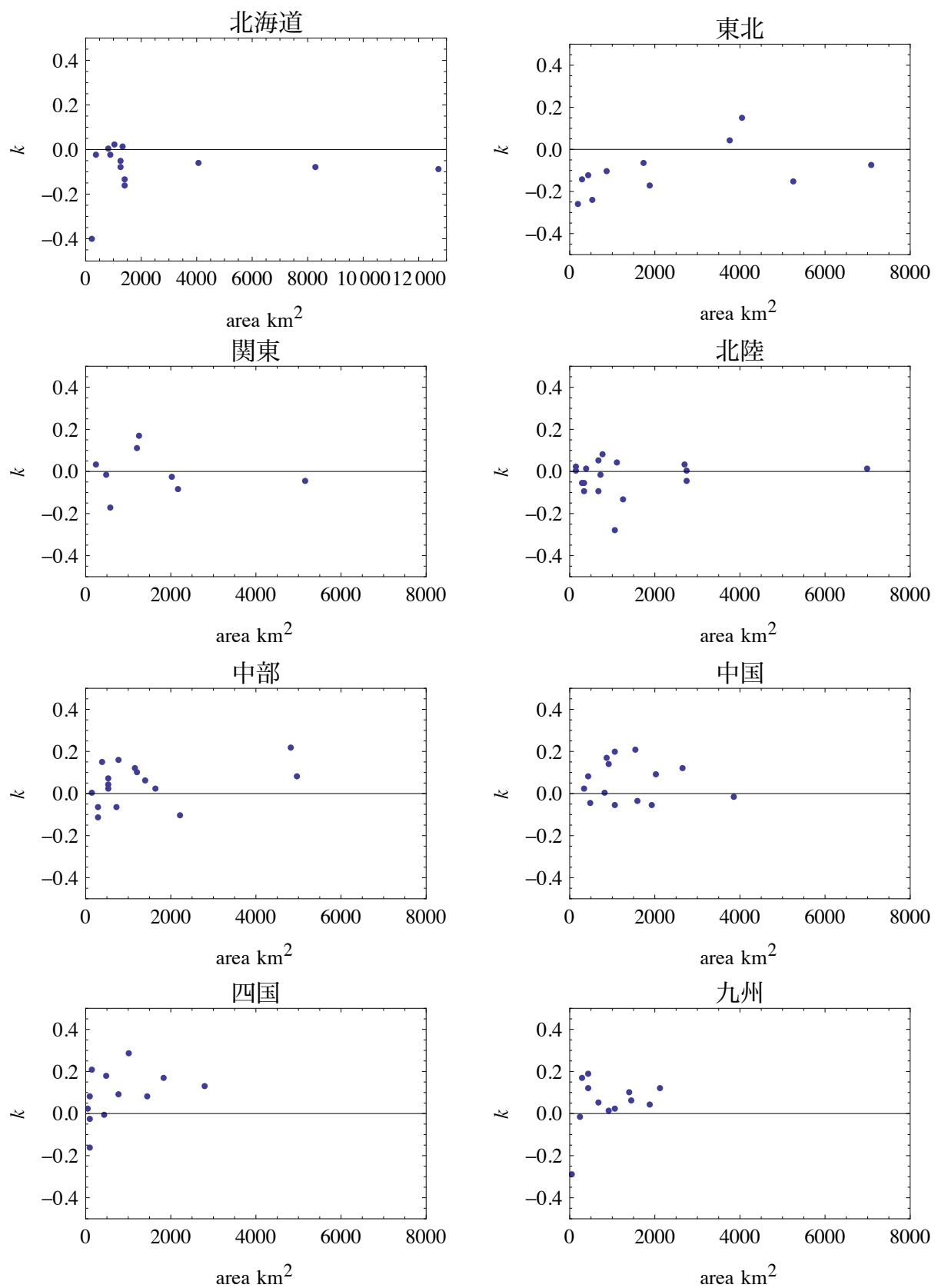


図4.7 流域面積とGEV分布の形状母数の関係

図4.8は第3章において示したL積率比 τ_3 と τ_4 の関係を各地域の3時間、6時間、12時間、24時間の流域平均雨量について求めたものである。すべてのデータは図3.1の中央右下のGumbel分布の τ_3 と τ_4 の点の付近にあるので、この部分のみ拡大して示している。この図を見ると、東北、関東および四国においてはほとんどのデータは一般パレート分布と一般ロジスティック分布に挟まれる範囲に分布しているが、他の地域ではこの範囲より外側のものが散見される。この範囲内にほとんどの分布が存在していることから、この範囲外にあるデータはこれまでに使われてきた分布モデルでは良好に適合させることが難しいことを示している。また、多くの地域で、AMS資料であるにもかかわらず、一般パレート分布をあてはめることが適当と考えられる資料が相当量存在していることも注目すべきである。

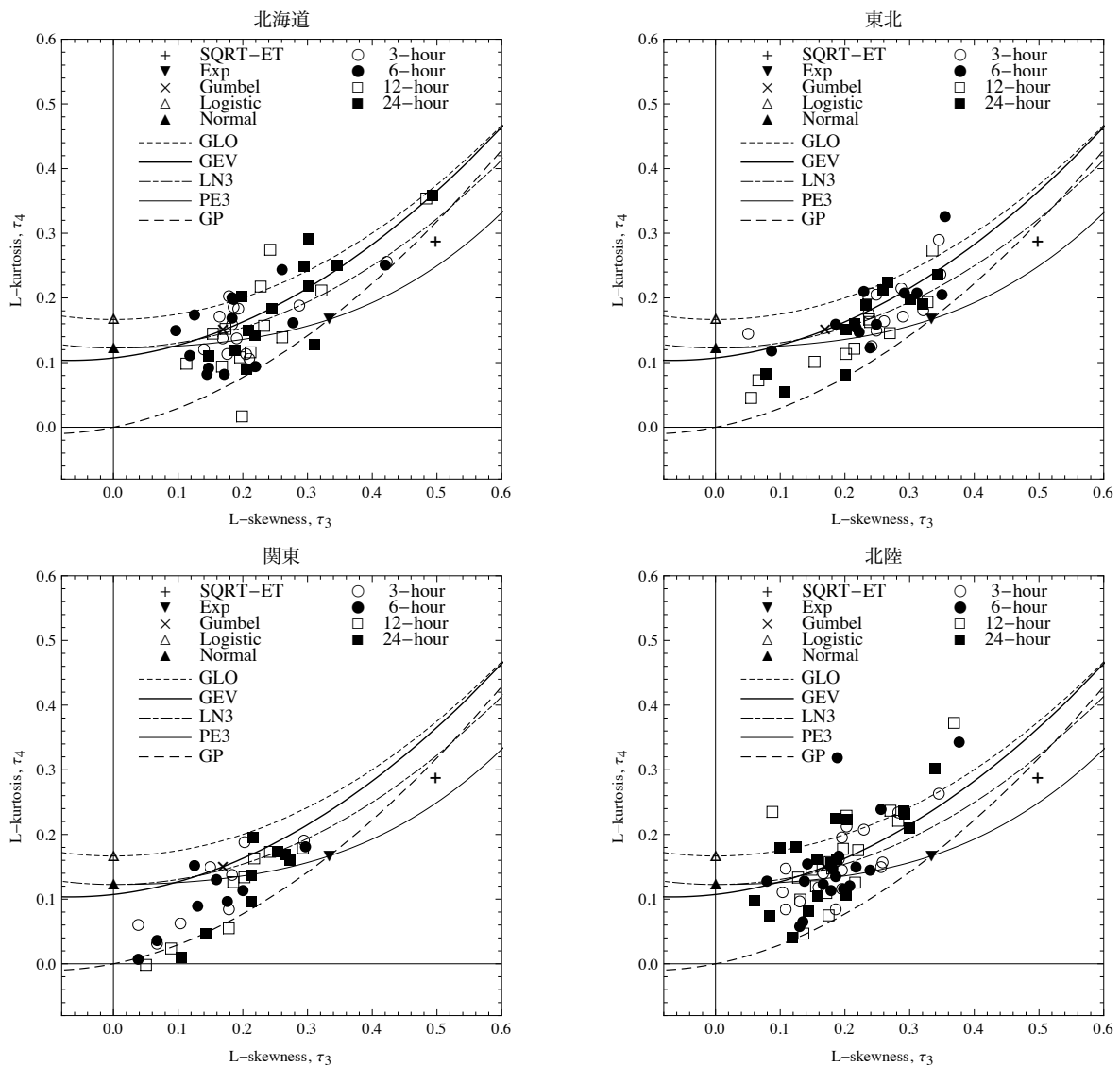


図4.8 3h,6h,12h,24hの短時間雨量毎の τ_3 と τ_4 の関係(1)

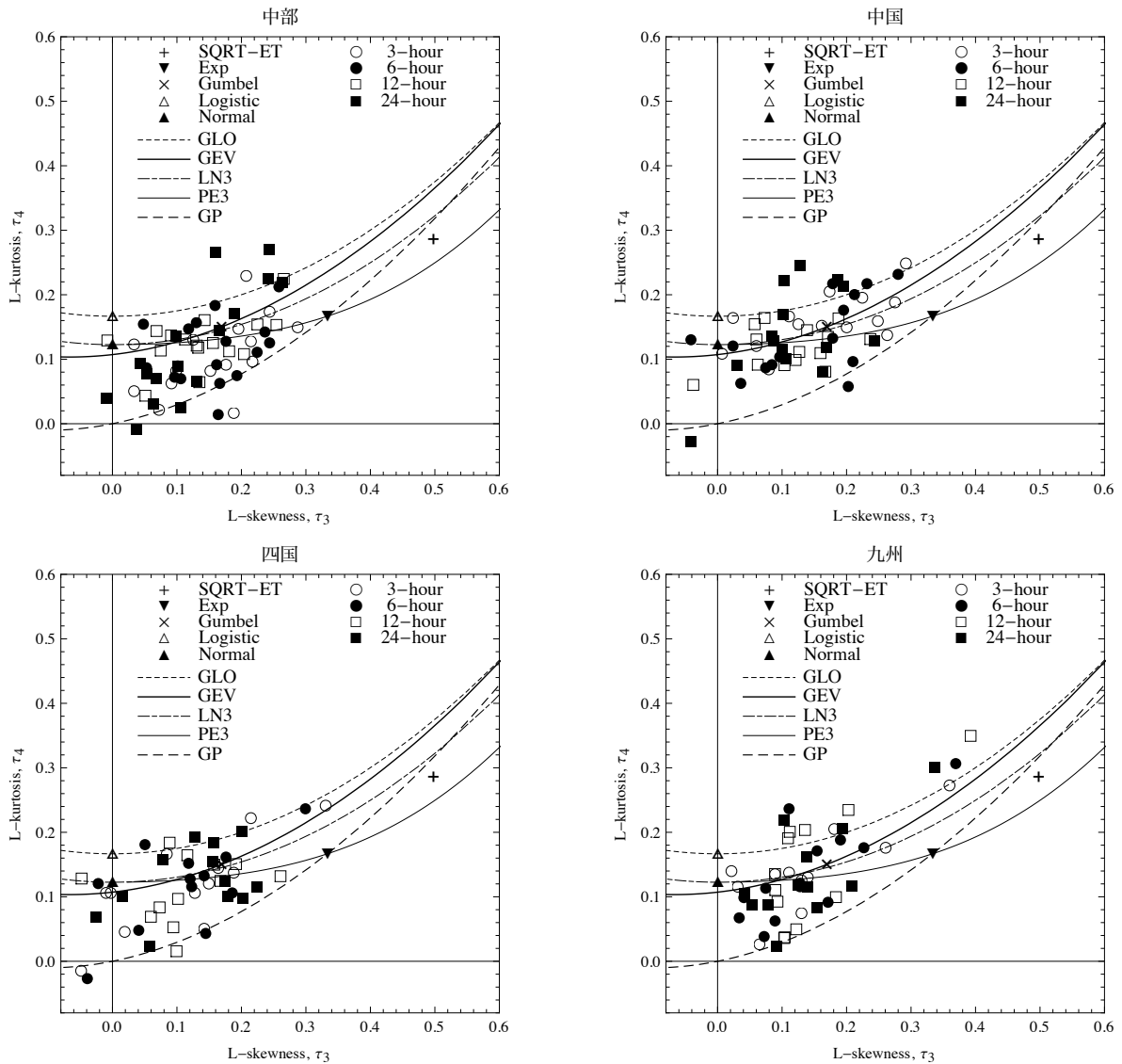


図4.8 3h,6h,12h,24h短時間雨量毎の τ_3 と τ_4 の関係(2)

図4.9はL積率比 τ_3 と τ_4 の関係が一般ロジスティック分布および一般パレート分布で挟まれる範囲から大きく逸脱している事例として北陸地方の梯川埴田地点の6時間降雨量とりあげ、これをGumbel確率紙にプロットし、あわせていくつかの確率分布モデルをあてはめたものである。この例では上位4位と5位の間に約24mm離れており、5位以下のデータの並びと4位以上のデータの並びとは傾向が異なっている。確率紙上でデータの並びが滑らかでないこのような事例を対象に計画規模を決定する必要がある場合には、確率分布モデルの説得性が乏しいと思われ、計画担当者は大いに悩むことになる。なお、平方根指数型最大値分布は他の分布と異なる特徴を有しており、この場合にも明らかに説得性がない。

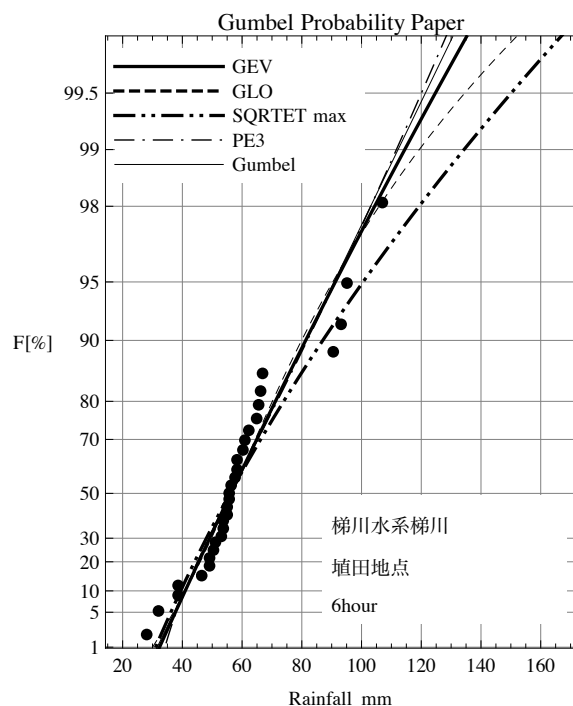


図4.9 梯川埴田地点上流平均6時間雨量の確率プロット

図4.8に示すようにこの標本のL積率比 τ_3 と τ_4 の関係は一般ロジスティック分布の曲線よりかなり上方に位置しており、どの分布をもってしても良好にあてはめることができない

4.2 洪水極値データの解析

次に洪水ピーク流量について適合度および安定性を検討する。比較に用いた資料は、わが国の68の一級水系における99箇所の流量である。観測流量は観測地点より上流の氾濫やダムなどによる貯留の影響を受けている。この影響を流出解析により取り除いた流量を解析の対象にした。流量の観測が始まった時期が異なるため、年最大値で整理されている99標本の観測年数は図4.10のような度数分布となる。40年～44年が40標本、35年～39年と45年～49年が18標本であり、平均値は42.1年である。

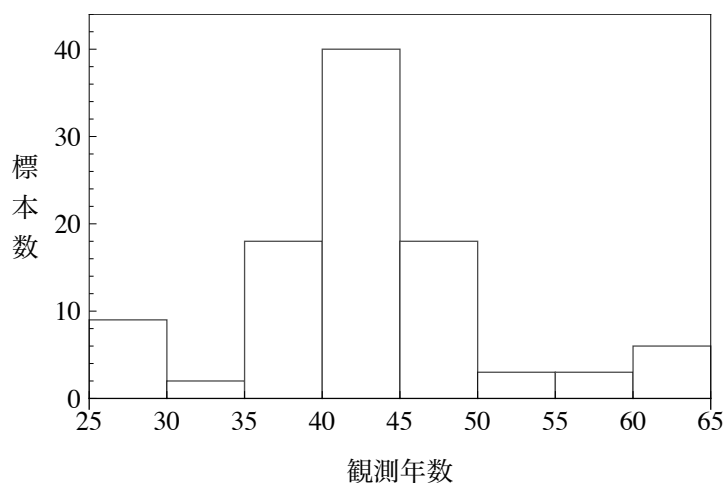


図4.10 用いた年最大値標本の観測年数

比較に用いた分布モデルは、指数分布【Exp】（L積率法）、Gumbel分布（L積率法）、平方根指数型最大値分布【SQRT-ET max】（最尤法）、一般極値分布【GEV】（L積率法）、対数ピアソンⅢ型分布（原標本の積率解【LP3Rs】、対数標本の積率解【LP3Ls】）、対数正規分布（3母数の場合岩井法【Iwai】、2母数の場合最尤法【LN2】）である。

なお、確率分布によっては標準変量の考え方がいくつもあり得るものがある。ここでは、平方根指数型最大値分布と一般極値分布の標準変量はGumbel分布と同じものを用いた。

また、SLSCの算定にあたっては、各種の分布に対して近似的に不偏なクォンタイルを与えるCunnaneの公式(Handbook of Hydrology, 1993)をプロットングポジションに用いた。

4.2.1 適合度評価

第3章に示したように適合度を測る規準としては相関係数、標準最小自乗規準(SLSC)、最大対数尤度、赤池の情報量規準(AIC)等がある。最大対数尤度とAICは、比較対象の確率分布の中での相対的評価しかできないので複数の標本に対して絶対的な数値基準を設定できないことおよび相関係数よりSLSCの方が規準の持つ意味が明確であることから、ここではSLSCにより適合度の評価を行うこととする。

図4.11に各確率分布毎に99の標本に対して求めたSLSCのヒストグラムを示す。なお、SLSCが0.07以上の階級には母数推定がうまく行かなかった場合や対数正規分布で対数変換標本値の歪み係数の絶対値が0.3を越えた場合を含めて表示している。この場合は対数正規分布をあてはめるのは不適切である可能性が高い（星，1998）。

SLSCが小さいものが多いのは、LP3Ls分布、岩井法、GEV分布である。これらはすべて3母数の確率分布モデルである。LP3Rs分布も3母数であるがSLSCが小さいものはそれほど多くはない^{注11}。一方、2母数の分布は、一般的に3母数の分布よりSLSCが小さいものが少なく、LN2分布のつぎにグンベル分布がSLSCが小さいものが多い。なお、LN2分布では全標本の44%で対数変換後の歪み係数の絶対値が0.3を越えている。

図4.12は各標本に対してSLSCが最小になった確率分布の度数を示したものである。LP3Ls分布が53%を占め、つぎにGEV分布、岩井法がつづく。LP3Rs分布とLN2分布は、今回対象とした標本ではSLSCが最小になることがなかった。

図4.13はある標本に対して8つの確率分布のうち最小のSLSCの度数分布である。SLSC=0.02以下が15%、0.03以下が67%、0.04以下が96%となっている。宝・高棹(1990)は、地点降雨資料を解析して、SLSC=0.02を十分な適合度と判断される基準、SLSC=0.03を満足すべき適合度の判定基準としている。しかしながら、今回対象とした流量資料では、満足すべき適合度の判定基準としてSLSC=0.03をとると、3分の1の標本は適合度を満足する確率分布モデルが1つもない状況になる。本来、適合度の評価は適合度の悪いモデルをふるい分けるものであり、結果としてどのモデルも残らないという状況になる判定基準は適当ではない。一方、図4.10に示したように観測期間が平均的に42年程度であることから、超過確率1/100～1/200の確率水文量は外挿により求めることになり、なんらかのモデルを必要とする。このような状況を踏まえ、河川流量の適合度の判定基準をSLSC=0.04程度に設定することが適当と考えられる（田中・宝，1999）。なお、SLSCは全体的な適合度を評価しており、標本

^{注11} LP3Rs分布は母数を求める際に原点まわりの1～3次のモーメントを用いるが、このモーメントの式が分母に $(1-a)^b$ 、 $(1-2a)^b$ 、 $(1-3a)^b$ を含んでおり、 a が正の場合、1、 $1/2$ 、 $1/3$ で不連続となる。また、星(1998)が示した近似式の適用範囲を逸脱するデータも多く存在する。星の近似式を使わずに母数を求めると、データの最大値より小さな上限値を有するような分布となることがあり、一見、相関係数でみる適合度は高いが、実際に生じた最大値をモデル上で再現できないという問題も含んでいる。

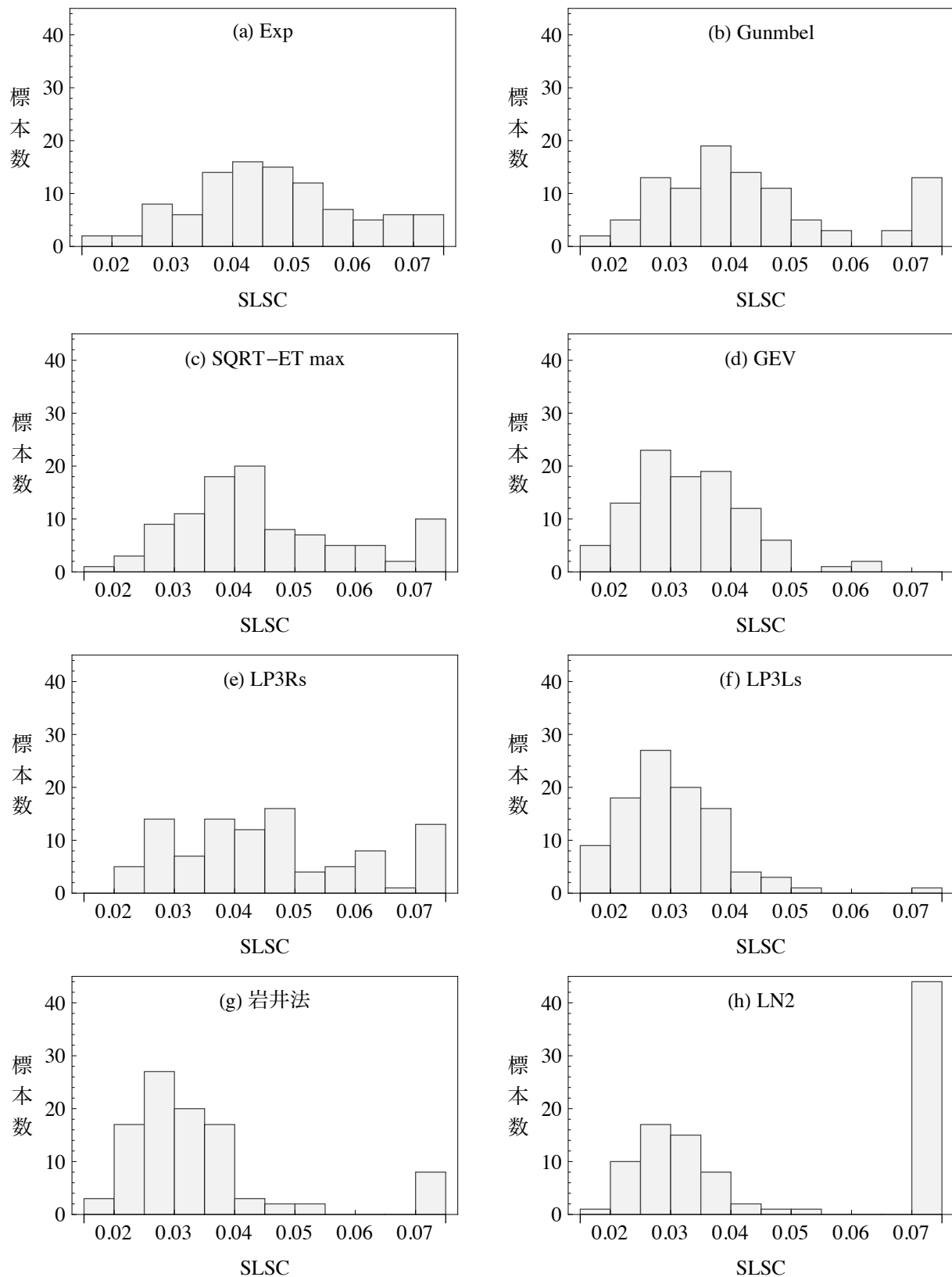


図4.11 各確率分布毎に99の標本に対して求めたSLSCのヒストグラム
SLSCが0.07以上の部分には母数推定がうまくいかなかった場合や対数変換値の
歪み係数の絶対値が0.3を超えているものを含んでいる

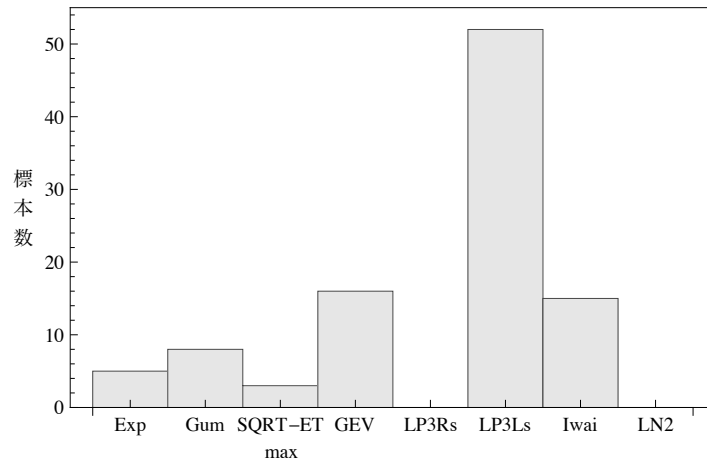


図4.12 各確率分布のSLSCが最小になった標本数

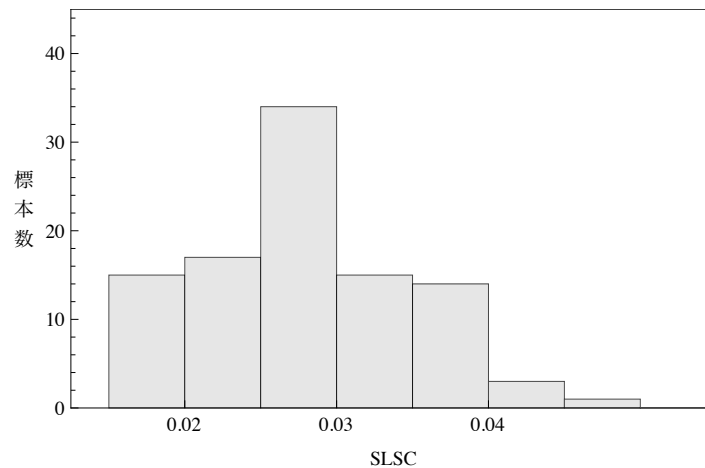


図4.13 確率分布の最小SLSCの度数分布

の最大値付近の適合度を示しているわけではない。頻度解析の目的は多くの場合、非超過確率の大きい領域におけるクォンタイルを求めることであり、この部分が確率紙上でデータとモデルの乖離が大きいと説得性が低下するので、SLSCでの数値的な判断だけでなく、確率紙上での確認をあわせて行う必要がある。

4.2.2 安定性評価

安定性の評価はリサンプリング手法で行う。リサンプリング手法にはjackknife法やbootstrap法がある。jackknife法は大きさ n の標本のうち任意の1データを欠いたデータ数 $n-1$ 個の標本を n セット作成し、これらの標本から求めた統計量をもとに不偏推定値およびそのまわりの推定誤差を算定する手法である。一方、bootstrap法は大きさ n の標本から重複を許して任意に n 個取り出した標本を複数作成し、これらの標本から求めた統計量をもとに不偏推定値およびそのまわりの推定誤差を算定する手法である。jackknife法は計算回数が少なく、作成する標本数および不偏推定値、推定誤差が一意的に定まるのに対し、bootstrap法は作成する標本数が任意に設定でき、作成する標本数によって不偏推定値や推定誤差が異なる。このため、今回の検討における確率分布モデル相互の安定性評価の指標としてはjackknife法を用いた。

図4.14は九州，緑川の洪水ピーク流量の年最大値についてLP3Rs分布とLP3Ls分布をあてはめるとともに，これらについてjackknife法を適用して得られたjackknife推定値と推定誤差を併せて示したものである．この観測所においては，低水路満杯流量が $2,700\text{m}^3/\text{s}$ であり，この付近の観測資料がなく，再現期間10年以上の範囲のバラツキが大きめである．しかしながら，SLSCは前者が0.023，後者が0.022であり，比較的よく適合している部類にはいる．実線はLP3Rs分布を示している．そのうち太い実線は全資料を用いた推定曲線である．3本の細い実線はjackknife法による推定曲線であり，中央の細線がjackknife推定値，その両側の細線が推定誤差の幅を示している．この両外側の曲線は観測値をほとんど包絡するように引かれている．また，中央の線と太い実線を比較すると，太い実線がjackknife推定値（中央の細線）より少し左に偏っており，同じ非超過確率に対し小さめの確率水文量を与えていることがわかる．

一方，破線はLP3Ls分布を示している．先程と同様に，細い破線の両側の線が観測値をほとんど包絡するように引かれているが，非超過確率の大きい部分でその広がりが大きくなっている．jackknife推定値と全資料を用いた確率水文量はあまり大きな差はないが，ほんの少しだけ全資料を用いた確率水文量がjackknife推定値を上回っている．

この場合，再現期間100年の確率水文量は，jackknife法により，LP3Rs分布が $+160\text{m}^3/\text{s}$ の修正，LP3Ls分布が $-30\text{m}^3/\text{s}$ の修正となり，お互いが近づく方向に修正されている．

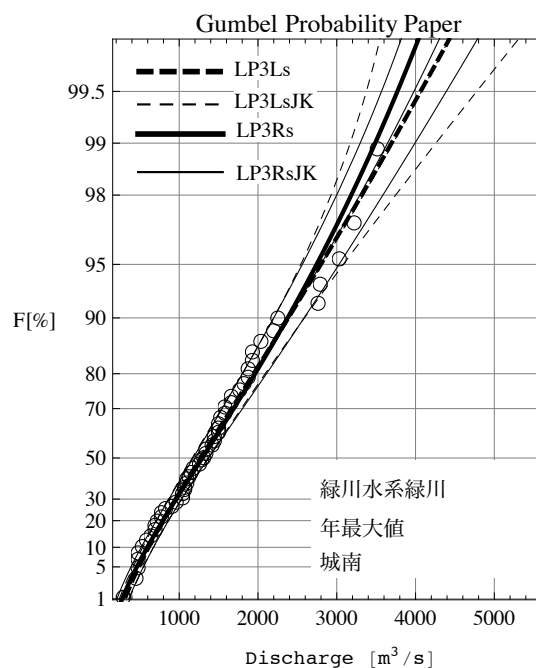


図4.14 年最大流量への対数PearsonⅢ型分布（LP3RsおよびLP3Ls）
のあてはめおよびそれらのjackknife推定値，推定誤差

表4.1は，緑川の洪水ピーク流量の年最大値について適合度と安定性の指標を示したものである．2母数対数正規分布は歪み係数の絶対値が0.3を越えたので不適(NG)としてある．また，岩井法はjackknife推定値が負になっており，適切に求められない．このような状況はStedingerら(1993)のクォンタイル法においてもいくつかの標本でみられることが確認された．この原因としては，これらの母数推定において，標本のすべてのデータが同じ寄与率を有していないことが上げられる．たとえば，岩井法では位置母数を決定する際に，標本の大きさを n とすると上位 $n/10$ 個のデータおよび下位

表4.1 適合度と安定性の指標（九州，緑川）

確率分布モデル	Exp	Gumbel	SQRT-ET max	GEV	LP3Rs	LP3Ls	Iwai	LN2
SLSC	0.042	0.023	0.033	0.023	0.023	0.022	0.022	NG
Q100全資料	4,260	3,703	4,663	3,750	3,487	3,746	3,887	NG
Q100jackknife推定値	4,260	3,703	4,681	3,736	3,646	3,717	-3,452	NG
Q100jackknife推定誤差率	9.50%	9.20%	9.60%	14.50%	9.50%	14.10%	-	NG
Q400jackknife推定誤差率	9.90%	9.60%	10.20%	19.60%	11.30%	18.80%	-	NG

の $n/10$ 個のデータが使われるため，jackknife標本において，上位および下位 $n/10$ 個のデータが同一のものに対する位置母数がほとんど同じになることによる．このような場合は，bootstrap法を適用する必要がある．SLSCでみると指数分布が適合度を満足しないと判断される．結局，5つのモデルが残り，これについて，jackknife推定誤差から判断すると，Gumbel分布が良さそうであるとともにAIC^{注12}でみてもGumbel分布が最小となり，選択すべきモデルとなっている．

図4.15は緑川の年最大値にGumbel分布をあてはめたものである．2母数であるので，推定誤差の幅が狭く，非超過確率の大きい部分でもそれほど広がることはない．なお，jackknife推定値が全資料を用いた推定曲線と重なっている．Gumbel分布や指数分布をLモーメント法で母数推定する場合にはjackknife推定値が全資料を用いた推定値と一致する．したがって，L積率法で母数推定するGumbel分布や指数分布は不偏推定値を与えるので適合度が満足されれば優れた分布モデルである．

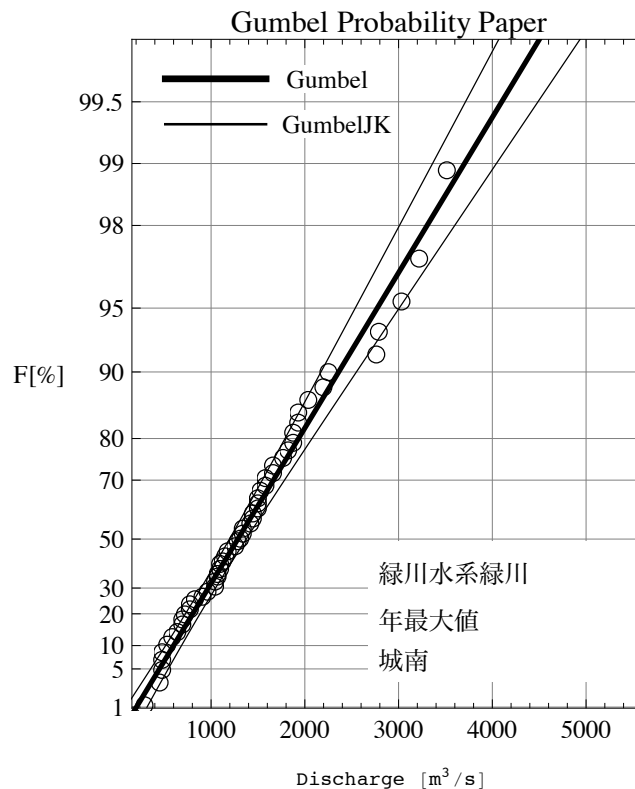


図4.15 年最大流量へのGumbel分布のあてはめおよびそれらのjackknife推定値，推定誤差

^{注12} AICは式(3.40)に示すように対数尤度を用いている．このため，本来は最尤法を用いて母数推定しなければならないが，ここではそれぞれの分布に最適な母数推定法によって求められた母数を用いて計算した擬似的なAICで判断した．

このように、リサンプリング手法を用いることにより、適合度が満足される確率分布モデルの中から安定性のよいモデルを選ぶことができる。その際、その結果を確率紙に表示して、あてはめたモデルとデータの関係を確認しておく必要がある。特に、3母数の分布モデルで上限値を持つようなあてはめ結果となる場合、なかには、最大値よりも小さいクォンタイルを与えることや上限値が上位のデータよりも小さいこともあるので上位のデータとあてはめたモデルとの関係を見ておくことが重要である。

4.3 結論

わが国の一級水系の基準点における3, 6, 12および24時間雨量の豪雨の解析から次のことが分かった。

- (1)GEV分布の形状母数 k による解析を行うことにより、北から南西部に向かうにつれ、 k の値が負から正に転ずるように分布している。
- (2)流域面積が小さい場合 k の値はばらついているが、大きくなるとばらつきが少なくなる。また、流域面積が大きくなると k がわずかではあるが大きくなる傾向がある。
- (3)L積率比による解析では、東北、関東および四国においてはほとんどのデータは一般パレート分布と一般ロジスティック分布に挟まれる範囲に分布している。他の地域ではこの範囲より外側のものが散見される。
- (4)多くの地域で、AMS資料であるにもかかわらず、一般パレート分布をあてはめることが適当と考えられる短時間降雨資料が相当量存在している

また、全国68の一級水系における99地点の年最大流量資料を統一的な手法で確率評価を行った。適合度と安定性から評価し、次のことがわかった。

- (1)各確率分布モデルの適合度を明らかにした。
- (2)河川流量の満足すべき適合度の基準として $SLSC \approx 0.04$ とするのが適当である。
- (3)安定性評価の指標としてjackknife推定誤差が有力な指標となる。
- (4)jackknife推定値、jackknife推定誤差を確率紙にプロットすることにより、選択する確率分布モデルの総合的なチェックが可能となる。

参考文献

- Stedinger, J.R., R.M. Vogel, and E. Foufoula-Georgiou: Frequency Analysis of Extreme Events, Chap. 18, Handbook of Hydrology, (Ed.) D. R. Maidment, McGraw-Hill, New York, pp. 18.1-18.66, 1993.
- 宝 馨：7.1.1 異常降雨記録，水文・水資源ハンドブック，水文・水資源学会編，朝倉書店，pp. 228-231, 1997.
- 宝 馨，高棹琢馬：水文頻度解析モデルの母数推定法の比較評価，水工学論文集，土木学会，第34巻，pp.7-12, 1990.
- 田中茂信，宝 馨：河川流量の頻度解析における適合度と安定性の評価，水工学論文集，第43巻，土木学会水理委員会，pp.127-132, 1999.
- 星 清：水文統計解析，開発土木研究所月報，No.540，pp.31-63, 1998.

5. 非毎年値資料を用いた水文頻度解析

従来、河川計画の基本量である外力を定める際には、毎年最大値資料(AMS, Annual Maximum Series)を対象とした水文頻度解析が主に用いられてきた。しかしながら、AMS資料の中には寡雨年の年最大流量等のように洪水とは呼べないようなものも少なからず含まれている場合がある。このような場合、AMS資料を用いて確率年100年から200年の確率水文量を求めるとその精度や信頼性について疑問が残る。

また、河川によっては第二次世界大戦前後において未曾有の洪水を経験しているものがあり、その被災経験が流域住民に強く残っている場合がある。このような場合、その大洪水の流量も含めて、その流域の洪水の発生頻度を推定すべきであるが、一般に戦中から戦後にかけての数年から十数年間は雨量や流量の記録が乏しく、従来のAMS資料を用いる方法が適用できない。

このような問題を回避できる方法として、時系列資料の中からある値より大きな水文量のみを抽出して標本とする非毎年値資料(PDS, Partial Duration SeriesまたはPOT, Peaks Over Threshold)による頻度解析がある。

本章では、POT資料を用いて確率水文量を評価する方法について紹介し、AMS資料との比較を行う。

5.1 非毎年値資料を用いることの意義

わが国の降雨および洪水ピーク流量の観測記録は約50年であるので、計画規模1/100～1/200の降雨量や洪水ピーク流量を求めるには外挿によらねばならない。第4章4.2.2安定性評価で示したようにjackknife推定誤差は推定値の非超過確率が大きくなるにつれて大きくなる。したがって、観測期間が長ければ、すなわち、標本サイズが大きければ標本の最大値の非超過確率が大きくなり、求めようとする計画規模に相当する非超過確率との差が小さくなり、推定誤差を小さく押さえることができる。過去に遡って観測をやり直すわけにはいかないので、現在蓄積されている資料を最大限利用しなければならぬ。POT資料はある閾値を超えるすべての独立した極値をデータとして扱うものであり、観測期間が短くても1年間にたくさんの事象が発生する場合には観測期間が標本サイズとなるAMS資料に比べ多くのデータを有する標本とすることができる。第3章で図3.11に示したように、AMS資料の最小値の非超過確率に相当するPOTデータの上位からの順位は、観測年数が40年の場合、観測年数の4倍余りである。したがって、年最大値において欠測のないように閾値を設定し観測を継続して得られたすべてのデータを蓄積すればAMS資料に比べ約4倍の大きさのPOT標本が得られるはずである。たとえば、図3.12は観測年数が25年であり、AMS資料に対しPOT資料は3.8倍であるが、確率紙上でAMS資料の最大値と最小値で挟まれる区間にAMS資料の3.8倍ものPOT資料が含まれるわけであるから、より精度の高い推定値が得られることが容易に理解される。従来の研究によると、POT資料は事象の年あたり発生率が1.65以上の場合、AMS資料に比べより正確な推定値を与える(Handbook of Hydrology,1993)。すなわち、POT資料の標本の大きさがそれを取り出した元の時系列資料から得られたAMS資料の標本の大きさに対して1.65倍以上であればPOT資料の方が正確な推定値を与える。

わが国の水文資料については、地点降雨資料は比較的整理、蓄積がしやすいため、時間雨量や日雨量の長期時系列記録があれば容易にこの条件を満たすことができるが、基準地点上流の流域平均雨量や洪水ピーク流量などは豪雨や洪水のイベントが少ない年があることやその資料整理に手間がかかることから大きな事象のみの間欠的な時系列資料として蓄積されている。

5.1.1 毎年度資料と非毎年度資料の関係

AMS資料とPOT資料の理論的關係およびその解析方法について、Handbook of Hydrology(1993)に詳しく記述されているのでそれを紹介する。

洪水、降雨、およびその他の多数の水文資料のモデル化には、一般的に2通りの方法が使用されている。AMS資料を用いる方法では、各年の最大の事象について検討する。部分的な水文資料(PDS)または閾値超過ピーク(POT, Peaks Over Threshold)とも呼ばれるPOT資料を使用する方法では、打切レベルまたは閾値レベルを超える全てのピークについて解析する。AMS資料を使用する方法では、各年の最大事象のみが用いられるため、ある年の2番目に大きな事象が他の年の最大事象を上回るか否かが考慮されないという問題がある。さらに、乾燥地域または準乾燥地域においては、渇水年における年最大洪水流量が0となる、またはあまりにも小さいため洪水と称するには無理がある場合も考えられる。

POT資料では、所定の閾値を上回るすべての独立なピークが対象とされるため、上記の問題は回避される。年超過確率は後述する式(5.4)を用いたPOT解析や経験的關係から推定することができる。POT解析の擁護者の主張によるとPOT記録は比較的長期的で信頼性の高い記録が使用できることが多く、閾値を超えるピークの発生率が十分に大きいとき（即ち、閾値を超過する確率が指数分布で表され、閾値を超える事象の発生間隔がポアソン分布にしたがう場合、1.65事象／年より大きいとき）POT解析を用いることによって、対応するAMS資料によるよりも高精度の確率水文量を推定できる。このような長所を持つ一方で、POT解析には、独立ピークを特定するための判定基準が必要になるという短所がある。したがって、POT解析は、AMS資料を使用する解析法に比して複雑な方法となる。

POT資料のモデル化においては二つの問題が生じる。一つは、閾値レベルを超える大規模事象の発生率をモデル化しなければならないという問題であり、もう一つはそれらの事象の大きさをモデル化することである。たとえば、事象の発生率のモデル化にはポアソン分布が使用されることが多く、閾値を上回るピークの規模を記述するためには、指数分布が使用される。非超過確率の大きい事象については、複数のモデルで1年当たりの平均発生回数が一致するのであれば、発生に関する確率モデルはそれほど重要ではない。

AMS資料とPOT資料における事象の頻度の間には、いくつかの一般的な関係が存在する。POT資料について、 λ を閾値 x_0 を上回る事象の1年当たりの平均発生回数とする。 $G(x)$ は、事象が x を下回る確率、したがって (x_0, x) の範囲をとる確率とする。 $x \geq x_0$ となる任意のレベル x について、発生率は次の式で表される。

$$\lambda^* = \lambda (1 - G(x)) \quad (5.1)$$

対応するAMS資料に関する累積分布関数 $F(x)$ は、ある年に関する年間最大値が x を超えない確率である。独立事象については、1年間を通じて x を超過しない確率はポアソン分布によって与えられ、次の式で表される。

$$F(x) = \exp(-\lambda^*) = \exp\{-\lambda (1 - G(x))\} \quad (5.2)$$

この関係は、1年を発生率が λ^*/m の m 個の区間に分割することによって導出される。 λ^*/m が小さいときは、1年間の非出現確率は基本的に $(1 - \lambda^*/m)^m$ となる。式(5.2)は、 $m \rightarrow \infty$ のときの極限において得られる。

式(5.2)から、AMS資料に関する累積分布関数と、POT資料に関する発生率および分布との関係が明らかになる。リターンピリオド T_a とすると年超過確率 $1-F(x)$ は $1/T_a$ で表される。POT資料においてこれに対応する超過確率 $(1-G(x))$ を q_e とすると、式(5.2)は、次のように表すことができる。

$$\frac{1}{T_a} = 1 - \exp(-\lambda q_e) = 1 - \exp\left(-\frac{1}{T_p}\right) \quad (5.3a)$$

ここに、 $T_p=1/\lambda q_e$ はPOT資料におけるレベル x に関するリターンピリオドである。 T_p について式(5.3a)を解くと、次の関係式が得られる。

$$T_p = -\frac{1}{\log(1 - 1/T_a)} \quad (5.3b)$$

POT資料においては1年間に起こりうる事象は1を上回るため、 T_p は T_a よりも小さい。

式(5.3a)によって、 x を上回る事象の平均発生率 λq_e は、AMS資料における年超過確率 $1/T_a$ に変換される。 $T_a > 10$ 、すなわち発生頻度の低い事象については、年超過確率 $1/T_a$ は、POT資料に関する平均発生率 $\lambda q_e = \lambda(1-G(x))$ にほぼ等しくなる。したがって、 $T_a = T_p$ となる。

式(5.2)において、閾値 x_0 を超える事象の規模を一般パレート(GP, Generalized Pareto)分布

$$G(x) = 1 - \left\{1 - k \left(\frac{x - x_0}{\alpha}\right)\right\}^{1/k} \quad k \neq 0 \quad (5.4a)$$

で表すことができるとすると、 k が正の値のとき、累積分布関数は上界 $x_{max} = x_0 + \alpha/k$ を持つ。 $k < 0$ のときは非有界であり、長く尾を引く分布となる。 $k=0$ のときは、式(5.4b)のような指数分布となる。

$$G(x) = 1 - \exp\left(-\frac{x - x_0}{\alpha}\right) \quad k = 0 \quad (5.4b)$$

式(5.4a)を式(5.2)に代入すると、 $k \neq 0$ のときに x_0 を上回るAMS資料に関する一般極値分布(GEV, Generalized Extreme Value)が得られる。

$$F(x) = \exp\left[-\left(1 - k \frac{x - \xi}{\alpha^*}\right)^{1/k}\right] \quad k \neq 0 \quad (5.5a)$$

$k = 0$ のときはGumbel分布となり、次の式で表される。

$$F(x) = \exp\left[-\exp\left(-\frac{x - \xi}{\alpha}\right)\right] \quad k = 0 \quad (5.5b)$$

ここに、 $x \geq x_0$ に対し変換後の母数 ξ および α^* は次のように定義される。

$$\xi = x_0 + \frac{\alpha(1 - \lambda^{-k})}{k}, \quad \alpha^* = \alpha \lambda^{-k} \quad \text{for } k \neq 0 \quad (5.6a)$$

$$\xi = x_0 + \alpha \log(\lambda) \quad \text{for } k = 0 \quad (5.6b)$$

この一般的なポアソン-パレートモデルは、柔軟性が大きく、多くの現象について妥当なモデルとなっている。AMS資料から得られたGEV分布の形状母数 k の地域推定値およびPOT資料から得られた一般パレート分布の形状母数 k の地域推定値は互換的に使用できる。

実際には発生率 λ は、1年当たりの x_0 を超過する平均回数という単純な方法で推定できる。そのほかのパラメーターは、次のようになる^{注13}。

$$k = \frac{\mu - x_0}{\lambda_2} - 2, \quad \alpha = (\mu - x_0)(1 + k) \quad \text{for } k \neq 0 \quad (5.7a)$$

$$\frac{1}{\beta} = \alpha = (\mu - x_0) \quad \text{for } k = 0 \quad (5.7b)$$

ここに、 $\mu = \lambda_1$ は標本 X の1次L積率（平均）、 λ_2 は2次L積率、 β は指数分布の尺度母数である。

以上、POT資料とAMS資料にあてはめられる分布として指数分布とGumbel分布または一般パレート分布とGEV分布の関係についてみてきたが、同じような説明が 星(1998)にもみられる。これらと同様の関係が平方根指数分布と平方根指数型最大値分布の間にも存在する（江藤ら、1986）。

第4章では、GEV分布を用いて全国のAMS資料を対象にした検討を行ったが、POT資料について考える際、指数分布は一般パレート分布の形状母数 k が0の場合に相当することおよび平方根指数分布に比べて一般パレート分布の方が図3.8で示したようにより広い範囲に適用でき、柔軟であることをふまえ、ここでは一般パレート分布の形状母数 k による解析を行う。なお、対象とした資料は第4章4.1でAMS資料を抽出した際の元の時系列資料から得られるPOT資料であるのでAMS資料とも特性を比較する。

また、洪水資料に対してもPOT資料の頻度解析を行い、AMS資料との比較を行う。この際、AMS資料を用いた確率水文量との比較を行うため、POT資料のうち毎年値と同じ個数だけデータを抽出する非毎年最大値を用いた頻度解析を行なうとともにその特性を比較検討する。

さらに、POT資料による解析の際の課題である閾値の選定方法に関する検討を流域平均雨量について行う。

5.2 豪雨のPOT資料の解析

各地方の各河川の流域平均雨量のPOT資料を指数確率紙にプロットすると降雨強度の小さい下位のデータがそれより上位のデータとは異なる特性をもっているものが多い。図5.1はその事例である。図5.2はこのAMS資料にGEV分布を当てはめた場合の形状母数 k とこのPOT資料にGP分布をあてはめた

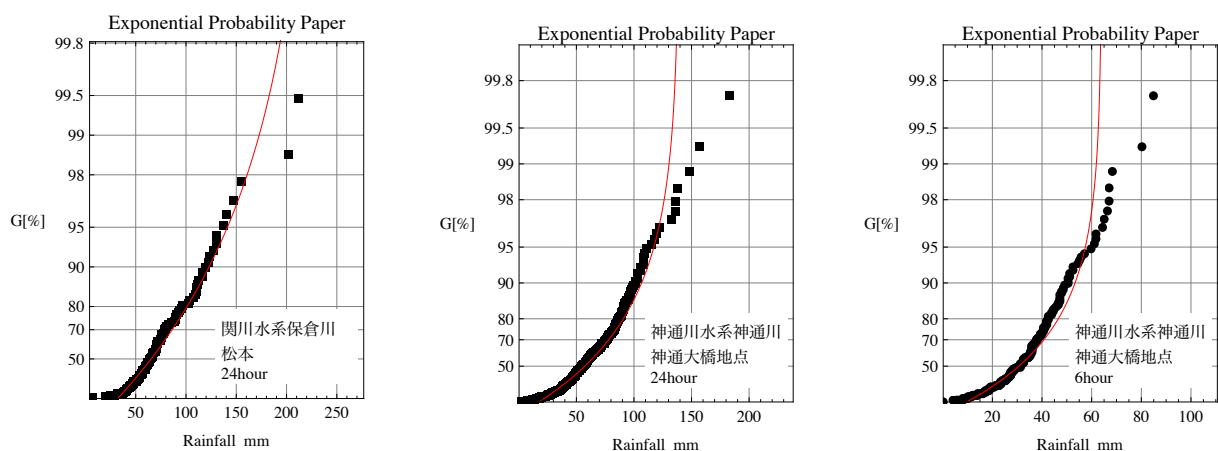


図5.1 下位のデータが確率分布モデルのあてはめに影響している事例

注13 田中・宝 (1999a)における k の式は-2が落れている。式(5.7a)が正しい。

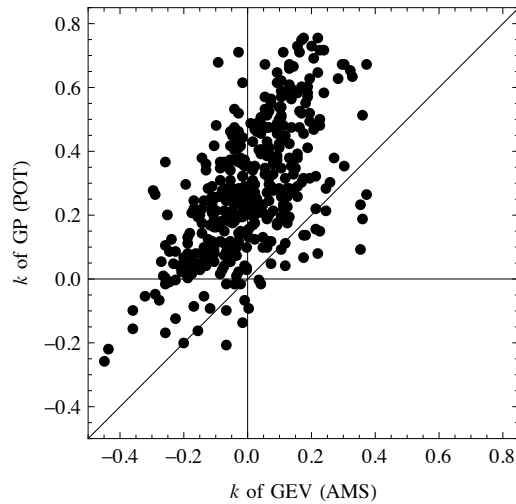


図5.2 AMS資料にGEV分布をあてはめた k とPOT資料にGP分布をあてはめた k の比較

場合の形状母数 k を比較したものである。両者の形状母数は互換的に使用できるものはきわめて少ないことが分かる。このようなデータの場合、このままモデルをあてはめると下位にある多くのデータがモデルの母数推定に支配的となり、上位のデータはモデルのラインから大きく外れてしまい、治水計画で対象とする超過確率の小さい部分についてクォンタイルを求めたり、超過確率を求める場合などにおいて説得性が乏しく、適切ではない。

POT資料の解析においては、本来、それぞれのデータの特性をしながら適切な閾値を設定すべきであるが、ここでは、まず、個々に適切な閾値を設定するのではなく、一律に各AMS資料の最小値を閾値とし、それ以上のデータをPOT資料として扱うものと、AMS資料と同じ個数だけ上位から取るものを考える。なお、それぞれにGP分布がどのようにあてはまるのかもあわせて検討する。

図5.3はAMS資料の最小値を閾値とした場合のPOT資料の観測期間 N_y と標本サイズ n の関係を示したものであり、図中に $n=N_y$ や $n=1.65N_y$ 、 $n=3.3N_y$ および $n=4.95N_y$ のラインをあわせて示してある。 $n=1.65N_y$ の線より左には15地点のデータが含まれている。図5.4はこのPOT資料にGP分布を当てはめた際の形状母数 k を図5.2と同様に示したものであり、図中の●は $n < 1.65N_y$ に対応している。POT資料で求めたGP分布の k がAMS資料で求めたGEV分布の k より大きめとなっており、なかでも●が全データのなかで上方に位置しているのが分かる。図5.4は図5.2に比べ少し対角線上に集まってきているがまだ多くのデータは対角線から上方にずれている。図5.5はAMS資料と同じ個数となるPOT資料につ

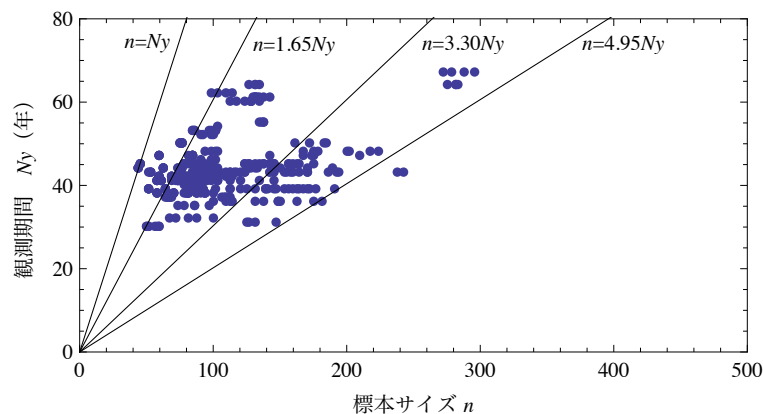


図5.3 AMS資料の最小値を閾値としたPOT資料の観測期間と標本サイズの関係

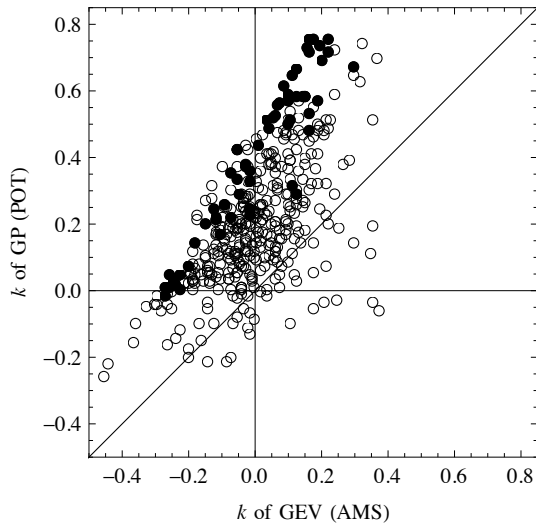


図5.4 AMS資料にGEV分布をあてはめた k とPOT資料にGP分布をあてはめた k の比較
(AMS資料の最小値を閾値とした場合)
●は図5.5で $n < 1.65N_y$ のデータに対応している

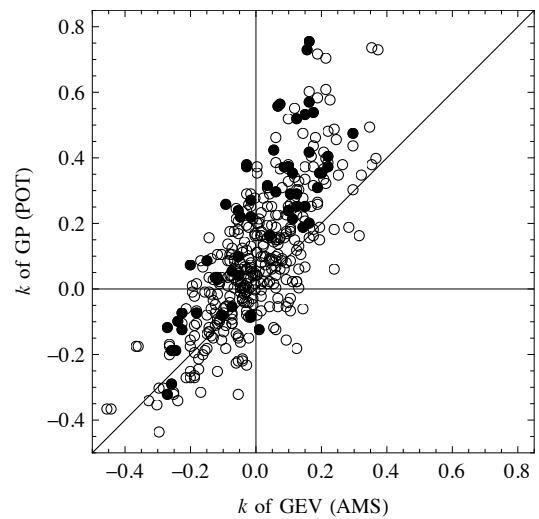


図5.5 AMS資料にGEV分布をあてはめた k とPOT資料にGP分布をあてはめた k の比較
(AMS資料と同じ個数上位から抽出した場合)
●は図5.4の●に対応している

いて同様に示したものである。図5.4に比べ●の分布が全データにある程度均等に分布するように改善されているのが分かる。また、図5.5ではかなりの数のデータが対角線上に乗るようになってきている。しかしどちらも対角線より少し強い傾きを有している。このような特性をふまえ、ここではPOT資料としてAMS資料と同じ個数だけ上位から取る非毎年年度最大値を考慮することとする。

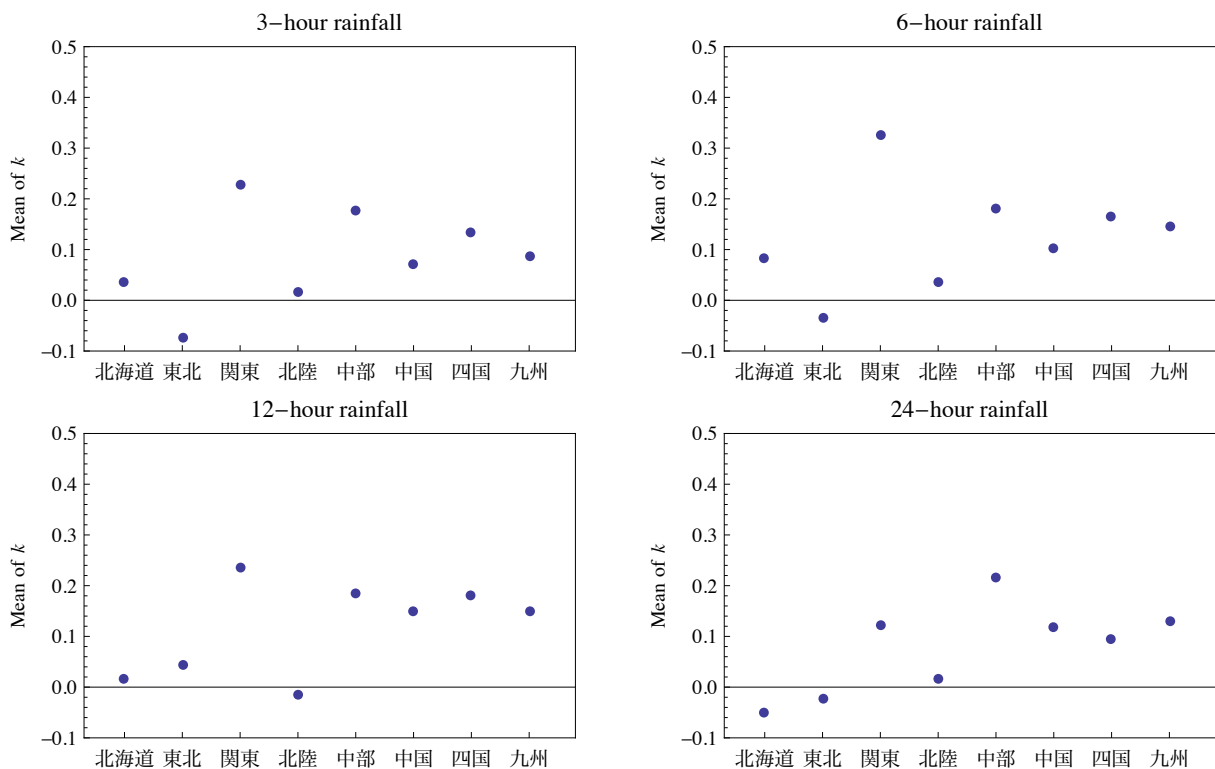


図5.6 3時間、6時間、12時間、24時間の流域平均雨量POT資料の k の地域内平均値

図5.6はそれぞれ3時間、6時間、12時間および24時間の流域平均雨量に対しGP分布をあてはめた際の形状母数 k の地域内平均値を示したものである。すべての時間区分においても北から南の地域に移るにつれ、 k の平均値が大きくなる傾向にある。

図5.7は図5.6の4つの時間区分のデータを平均したものであり、図4.3に示したAMS資料に対して求めたGEV分布の形状母数 k の平均値もあわせて示している。この図より、AMS資料の k よりPOT資料の k の方がどの地域においても大きくなっており、特に関東、中部ではかなり大きく、次いで北海道、東北が大きい。形状母数 k が大きくなることはあてはめた分布が右に長く尾を引く分布から上限値を持つ分布にむかって変化することを示している。

形状母数が流域面積の大きさとどのような関係があるか調べてみたが、AMS資料の場合とほぼ同じ結果であった。

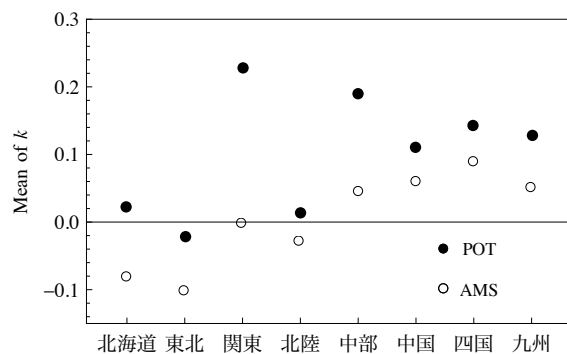


図5.7 流域平均雨量の k の地域内平均値の4つの時間区分の平均値

黒丸はPOT資料、白丸はAMS資料に対応している。どの地域においてもPOT資料の方がAMS資料より k の値は大きくなっている。

図5.8aはPOT資料のL積率比 τ_3 と τ_4 の関係を示したものであり、点線で τ_3 と τ_4 の重心を示した。また、これと比較できるようにAMS資料に対するL積率比も同様に図5.8bに示した。POT資料のプロットの多くはGP分布の τ_3 と τ_4 の関係に近く、一部一般ロジスティック分布の曲線に近いものも見られる。また、GP分布の曲線より下方に位置するプロットも少なからず見られる。これに対し、AMS資料で

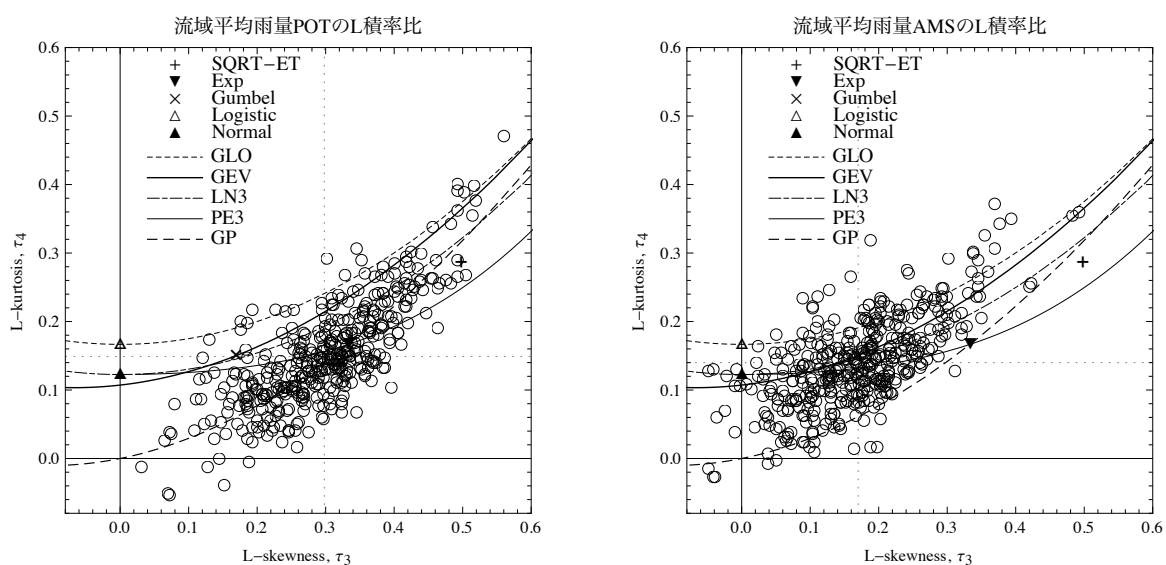


図5.8 流域平均雨量のL積率比の τ_3 と τ_4 の関係 (a: POT資料, b: AMS資料)

は全体的にGEV分布の曲線に沿ってプロットされているが、ばらつきの範囲が広く、GP分布から一般ロジスティック分布でカバーされる範囲より広い範囲に分布している。3母数対数正規分布(LN3)はGEV分布およびGP分布を横切るような形の曲線であるため、多くの資料に対し、LN3分布をあてはめることが可能となっている。プロットの重心はPOT資料については $(\tau_3, \tau_4) = (0.298, 0.149)$ であり、2母数の分布では指数分布に最も近い。AMS資料の重心は $(0.170, 0.140)$ であり、2母数の分布ではGumbel分布に最も近い。したがって、上述の関係より、概ねPOT資料は指数分布やGP分布であてはめられ、AMS資料はGumbel分布やGEV分布であてはめられることがわかる。なお、POT資料についてはロジスティック分布や正規分布をあてはめることが適当な資料はないが、AMS資料については少なからず存在するとともに τ_3 が負のものもみられる。

5.3 洪水ピーク流量のPOT資料の解析

本節では、洪水ピーク流量についてPOT資料として抽出したものとAMS資料として抽出したものを比較する。前節の結果をふまえ、POT資料は非毎年最大値（POT資料のうち統計年数と同じ数だけ上位から取り出したもの）とする。AMS資料とPOT資料の両面から適合度と安定性の指標を用いて評価し、両者を比較する。理論的には、年最大値と非毎年最大値は再現期間が10年より大きい範囲で一致する。前述したように、われわれが利用できる資料の観測期間は40年前後が多く、資料は多少にかかわらずバラツキを持っている。また、梅雨や台風による洪水がない年の最大流量は洪水と見なせないような値である場合がある。AMS資料を用いて求めた確率水文量とPOT資料を用いて求めた確率水文量は必ずしも一致しない。このような場合、AMS資料のみを用いて確率水文量を推定するより、POT資料からも確率水文量を推定し、両者を比較するほうがよりよい確率水文量の推定が行えると考えられる（田中・宝，1999a）

確率水文量の値やその変動特性は、確率分布モデルだけでなく用いる母数推定法によっても異なる。本研究で検討対象とした確率分布モデルと母数推定法は、AMS資料については、第4章と同様に、水文統計の分野で用いられてきた8種類の確率分布モデル（指数分布、グンベル分布、平方根指数型最大値分布(SQRT-ET max)、GEV分布、対数ピアソンⅢ型分布（原標本の積率解(LP3Rs)、対数標本の積率解(LP3Ls)）、対数正規分布（岩井法、2母数(LN2)の場合最尤法））を用いた。POT資料については、GP分布と指数分布（角屋の方法(角屋，1964)）、GP分布で形状母数 $k=0$ としたもの(GPExp))を用いた。GP分布の母数推定にはLモーメント法を用いた(Handbok of Hydrology, 1993)。

5.3.1 解析に用いた資料

1) 標本の観測年数

ここで解析に用いた資料はわが国の一級河川における79箇所での流量である。流量の観測が始まった時期が異なるため、標本の統計年数は、最大が62年、最小が27年、平均値が42.9年であり、40年～44年の階級が35標本あり、全体の約半数をしめる。

2) AMS資料とPOT資料の関係

POT資料には年間第2位や第3位等の洪水流量も含まれる。今回の検討では非毎年最大値を対象としているので、標本の大きさはAMS資料と同じである。もし、1年に1回しか洪水が発生せず、これ以外の期間は極めて小さい流量（AMS資料の最小値より小さい流量）だとすると、AMS資料とPOT資料は一致する。しかしながら毎年発生する洪水の年最大値のみが大きくて他の洪水は他のどの

年の年最大値よりも小さいものしか発生しないということはめずらしく、大洪水のあった年の第2位の洪水が他の年の年最大値より大きいことは十分考えられる。すなわち、AMS資料とPOT資料は上位から mc 個までは一致するがそれ以降は異なることとなる。一方、下位データに注目すると、POT資料の最小値はAMS資料の小さい方から me 番目よりも大きくなる。 mc と標本の大きさ n の比を上位一致率、 me と n の比を下位入替率と呼ぶことにすると、それらは図5.9aのようになる。大きな洪水の年間の発生回数が少ない河川では、上位一致率は高くなり結果として下位入替率が低くなる傾向がみとれる。一方、上位一致率が小さい河川では下位入替率が大きくなる傾向にある。全国のデータで見ると最も低いデータを除き、下位入替率は2割から4割余りとなっている。図5.9bはデータの順位の代わりに変数の大きさをもって示したもので、上位一致区間率は、 $(\text{AMS資料の最大値} - mc\text{番目洪水ピーク流量}) / (\text{AMS資料の最大値} - \text{AMS資料の最小値})$ であり、下位入替区間率は、 $(\text{POT資料の最小値} - \text{AMS資料の最小値}) / (\text{AMS資料の最大値} - \text{AMS資料の最小値})$ である。順位でみた上位一致率に比べ変数の大きさでみた上位一致区間率の方が大きな値となっている。これは既往のデータにおける超過確率の小さい洪水の記録が比較的頻繁に起こる中小規模の洪水に比べ右裾にまばらに分布していることを示している。右裾の分布の程度はGEV分布の形状母数に対応して変化する。この状況を見るため、すべてのAMS標本についてGEV分布の形状母数 k を求めた。図5.10はこの k の度数分布であり、 k の最小値は -0.509 、最大値は 0.371 である。図5.11は上位一致率と上位一致区間率を比較したものであり、あわせてAMS資料のデータが形状母数 $k=0.3$ 、 0 （Gumbel分布）および -0.5 のGEV分布の曲線に乗っている場合の上位一致率と上位一致区間率の関係を示している。この図より、上位一致率と上位一致区間率はGEV分布の形状母数と良好な関係にあることが分かる。図5.12はAMS資料

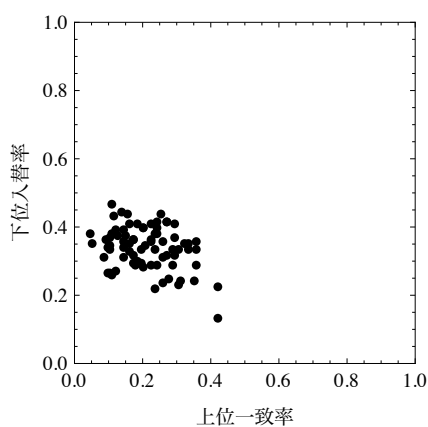


図5.9a AMS資料とPOT資料の関係
(標本内の順位で評価)

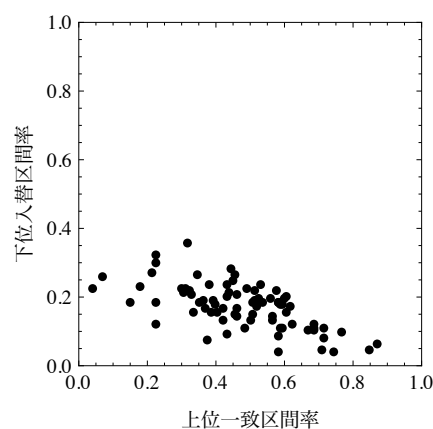


図5.9b AMS資料とPOT資料の関係
(変数の大きさに評価)

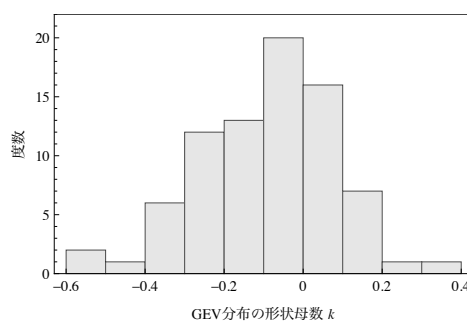


図5.10 洪水ピーク流量のAMS資料のGEV分布の形状母数 k の度数分布

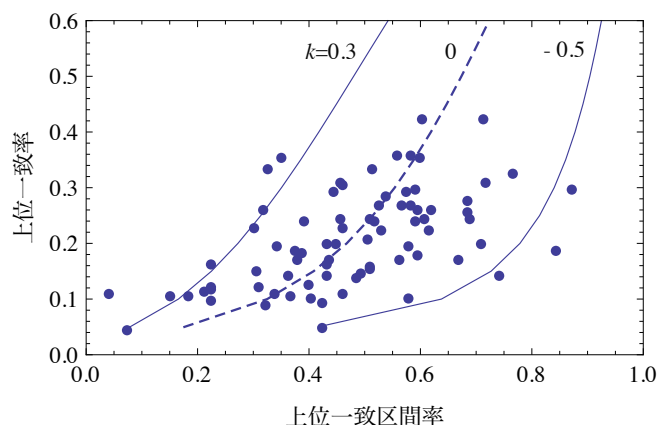


図5.11 上位一致率と上位一致区間率の関係
GEV分布の形状母数で上位一致率と上位一致区間率の関係が概ねわかる

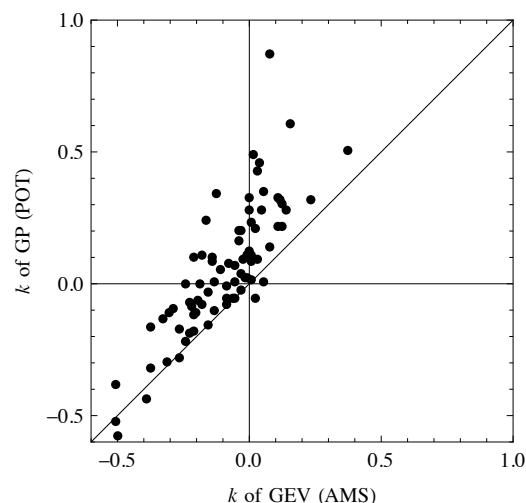


図5.12 洪水ピーク流量のAMS資料のGEV分布の形状母数 k とPOT資料のGP分布の形状母数 k の比較

のGEV分布の形状母数 k とPOT資料のGP分布の形状母数 k の比較を示している。POT資料およびAMS資料の k の平均値はそれぞれ0.057および-0.091であり、POT資料の k の方がAMS資料より少し大きい値を与える傾向にある。すなわち、POT資料はAMS資料に比べ右裾が短く上限を持ちやすい傾向にあることが分かる。ここで使用しているPOT資料とAMS資料は同じ観測期間のデータであり、上位のデータは同一であるので、 k の値が異なる原因は主に下位のデータの特性がPOT資料とAMS資料で異なることによるものであることが考えられる。 k が異なると特に外挿によって求める範囲のクォンタイルが異なることになる。これにかかわる影響の検討については第6章で述べる。

5.3.2 確率分布モデルの適合度

1) SLSCの算定

確率分布モデルによっては標準変量の考え方がいくつもあり得るものがある。ここでは、平方根指数型最大値分布とGEV分布の s^* はグンベル分布の s^* と同じものを用いた。

また、SLSCの算定にあたっては、プロットイングポジションを与える必要があるが、ここでは各種の分布に対して近似的に不偏なクォンタイルを与えるCunnaneの公式(Handbok of Hydrology, 1993)を用いた。

2) SLSCの算定結果と評価

図5.13aはAMS資料について最小のSLSCとなった確率分布モデルのヒストグラムを示したものである。LP3Lsが44標本で過半数を占める。次に適合度がよいのはGEV分布であり、これにグンベル分布、岩井法が続く。

POT資料について同様に整理したものが図5.13bであり、GP分布が最も多く43標本であった。この結果を見るとどちらの場合も3母数の確率分布モデルがデータへの適合という面では有利であることを示している。

図5.14は横軸および縦軸に、それぞれAMS資料の確率分布モデルの中での最小のSLSC値およびPOT資料における確率分布モデル中最小のSLSC値をとって比較したものである。AMS資料のSLSCはほぼ河川流量に対する良好な適合度の判定基準である0.04以内(田中・宝, 1999b)に収まっているがPOT資料の方は0.04で区切ると26標本において適合度が満足できる確率分布モデルを失ってしま

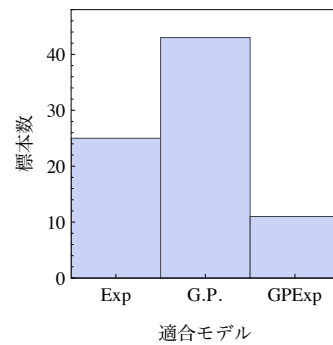
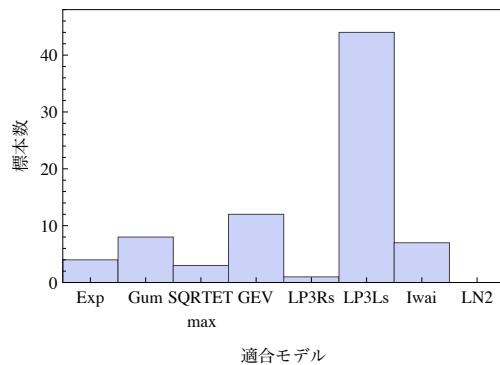


図5.13a AMS資料について最小のSLSCとなった 確率分布モデルのヒストグラム

図5.13b POT資料について最小のSLSCとなった 確率分布モデルのヒストグラム

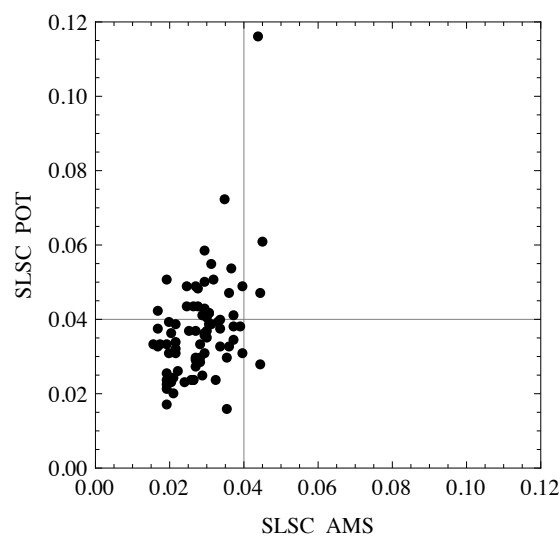


図5.14 AMS資料のSLSCとPOT資料のSLSCの比較

う、全体的には、少し正の相関があるが、強くはない。この図にプロットされていないが、SLSCが最小ではなく0.04より小さいものはたくさんあるので、AMS資料とPOT資料のどちらかで見のではなく、両者を見た上で判断することが望ましいといえる。また、ここでは、POT資料にAMS資料と同じ標本の大きさとなるよう上位データを用いたが、図5.1で示したようにPOT資料の下位データが上位データと異なる特性を有していることもあり得るので、閾値の設定も適切に行う必要がある。

3) jackknife法による確率水文量の不偏推定と毎年値資料と非毎年値資料の比較評価

確率分布モデルの安定性の評価はリサンプリング手法で行う。リサンプリング手法にはjackknife法やbootstrap法がある。jackknife法は大きさ n 個の標本のうち任意の1データを欠いた大きさ $n-1$ 個の標本を n セット作成し、これらの標本から求めた確率水文量をもとに不偏推定値およびそのまわりの推定誤差を算定する手法である。一方、bootstrap法は大きさ n 個の標本から重複を許して任意に n 個取り出した標本を複数作成し、これらの標本から求めた確率水文量をもとに不偏推定値およびそのまわりの推定誤差を算定する手法である。jackknife法は計算回数が少なく、作成する標本数および不偏推定値、推定誤差が一意的に定まるのに対し、bootstrap法は作成する標本数が任意に設定でき、作成する標本数によって不偏推定値や推定誤差が異なる。このため、今回の検討における確率分布モデル相互の安定性評価の指標としてはjackknife法を用いた。安定性の評価は確率水文量の推定誤差の大ききで判断した。

図5.15は渡川水系四万十川について、適合度のよいGumbel分布(SLSC=0.018, 実線)とGPExp分布(SLSC=0.021, 破線)のjackknife法による推定曲線であり、中央の線がjackknife推定値、両外側の線はjackknife推定値の両側に推定誤差の幅を示している。この両外側の曲線は観測値をほとんど包絡するように引かれている。また、毎年値資料から推定される確率水文量と非毎年値資料から推定される確率水文量が非常によく一致していることがわかる。

図5.16は仁淀川について適合度のよいLP3Rs分布(SLSC=0.018)とGPExp分布(SLSC=0.031)をあてはめたものである。AMS資料を用いた推定曲線(図中LP3Rs)は非超過確率0.95(再現期間20年)より大きい範囲でPOT資料による推定値と乖離している。しかし、jackknife推定値(LP3RSJK)はPOT資料による推定値とほぼ一致している。このようにjackknife法により確率水文量の偏りが補正されることがわかる。

また、四万十川と仁淀川の低水路満杯流量はそれぞれ $3,500\text{m}^3/\text{s}$ および $2,500\text{m}^3/\text{s}$ であり、POT資料のデータは、これより大きい範囲のみにある。したがって、POT資料のデータはすべて高水敷を流れる規模の洪水となっている。一方、AMS資料の方は低水路内だけで流れる出水も含まれていることになる。また、四万十川の毎年値資料には $3,000\text{m}^3/\text{s} \sim 3,500\text{m}^3/\text{s}$ 付近のデータが少なく、この周辺でデータが階段状にプロットされ、これが適合度を下げる一因にもなるが、四万十川の場合、非毎年値資料ではこういったことも回避できる。

このように、リサンプリング手法およびAMS資料とPOT資料を併せて評価することにより、偏りを補正した確率水文量の推定が可能となるばかりでなく、AMS資料固有の問題も回避でき、確率水文量の信頼性を向上させることができる。また、はじめに述べたように、AMS資料が適用できない標本に対して、残された手法はPOT資料である。ここで用いた手法は、そのような場合に対しても確率水文量を精度よく推定する有力な手段となる。

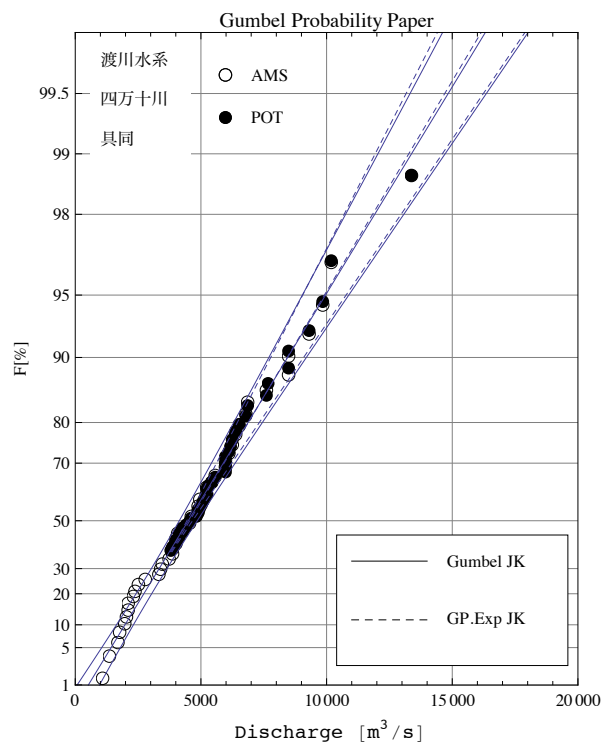


図5.15 AMS資料とPOT資料にそれぞれGumbel分布とGPExp分布をあてはめたもののjackknife推定値、推定誤差

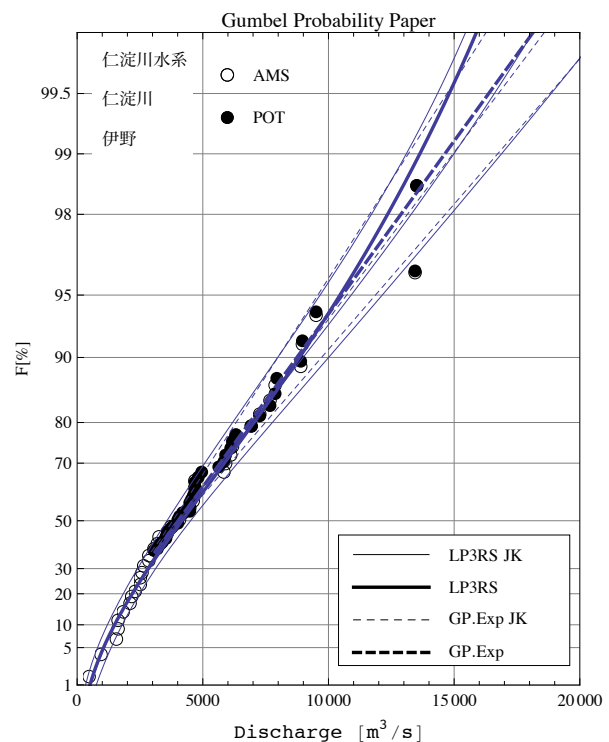


図5.16 AMS資料とPOT資料にそれぞれLP3RS分布とGPExp分布をあてはめたもののjackknife推定値、推定誤差

5.4 POT資料における閾値の選定方法

わが国では従来治水計画の立案にあたって使用されてきた頻度解析の手法はAMS資料を用いる方法であった。大河川の水文資料は戦前から蓄積されているが、多くの河川では戦時中から戦後しばらくの間観測がなされておらず、1950年代から継続した観測が行われ、資料が蓄積されてきている事例が多い。しかしながら、戦前に未曾有の洪水を経験した河川もあり、このような河川ではその大洪水の記録を治水計画に活かさずに比較的小さい規模の治水計画を提案したのでは流域および沿川住民の理解を得ることは難しい。このような場合、AMS資料では適切に頻度解析を行うことができないが、POT資料として整理すれば頻度解析が可能になる。また、観測期間が短い場合にはAMS資料に比べPOT資料を用いた頻度解析の方がよいことが指摘されている(Rasmussen et al., 1994)。

これまでAMS資料の解析に対してPOT資料の方が優れている点や柔軟な点を紹介してきたが、POT資料の解析では閾値の設定の課題がある。これまで、POT資料を扱う時にAMS資料の最小値を閾値とする場合やAMS資料と標本サイズが同じになるように閾値を設定し非毎年年数最大値とするなどの方法でPOT資料の解析を行ってきた。しかしながら、第3章図3.9や本章図5.1に見られるように閾値の設定如何によってあてはめられる分布やそのパラメーターが変化し、どのように設定するのが適当かという問題がある。洪水ピーク流量ならたとえば高水敷を流れる規模以上の洪水を対象にする等であれば物理的な意味があるが、地点雨量や流域平均雨量等のように物理的な閾値の設定根拠の無い場合の閾値の設定は難しい。

Tanaka & Takara(2002)は6種類の指標を用いて、POT資料を用いて解析する場合における閾値の選定方法（上位何個のデータを用いるかの選定方法）を比較している。本節では、これに基づいて検討する。

5.4.1 検討に使用した資料、確率分布モデルおよび母数推定法

検討に用いた資料は東北地方の最上川の下野地点(上流流域面積3,534km²)と両羽橋地点(同6,519km²)における流域平均雨量である。観測期間は1955年から1996年($N_y=42$)であり、この期間内の出水をもたらした上位129降雨が抽出された。降雨継続時間は3, 6, 12, 24時間である。

検討に用いた分布モデルは指数分布とGP分布であり、閾値 x_0 （位置母数）を超える確率 $G(x)$ はそれぞれ次のように表される。(式5.4の再掲)

$$G(x) = 1 - \exp\left(-\frac{x-x_0}{\alpha}\right) \quad k=0 \quad (5.8a)$$

$$G(x) = 1 - \left\{1 - k \left(\frac{x-x_0}{\alpha}\right)\right\}^{1/k} \quad k \neq 0 \quad (5.8b)$$

ここに α と k はそれぞれ尺度母数および形状母数である。

母数推定はL積率法(たとえばHandbook of Hydrology, 1993)を用いた。位置母数 x_0 として上位 r 個の最小値をとった指数分布およびGP分布をそれぞれ“Exp”および“GP”と表記することとする。また、最小値を用いずに高次のL積率を用いて位置母数を推定したものをそれぞれ“Exp1”および“GP1”と表記する。

5.4.2 閾値選定指標の有効性

流域平均雨量の閾値を選定する際に規準となりうる物理的な閾値はないので、以下の6つの指標について、 r を変化させて最も適切と考えられる r を求めてみた。

- (1)超過確率1/100に相当するクォンタイル (図5.17)
- (2)分布モデルの形状母数および尺度母数 (図5.18および図5.19)
- (3)SLSC (図5.20)
- (4)jackknife推定誤差 (図5.21)
- (5)形状母数を用いた自動選択指標 (図5.22および図5.23)
- (6)標本平均超過関数(Sample mean excess function, SMEF) (図5.24)

なお、後述するように下野地点と両羽橋地点双方の3, 6, 12, 24時間の資料について検討したが、ここでは下野地点の3時間の流域平均雨量の結果のみを示す。

1)超過確率1/100に相当するクォンタイル

図5.17は上位 r 個の資料について指数分布およびGP分布の超過確率1/100のクォンタイルを2種類の方法で推定した母数を用いて求めたものを示したものである。クォンタイルでの判断は、値の安定性だけでなく、2母数の分布と3母数の分布での値の傾向もあわせてみる。Exp分布およびGP分布のクォンタイルは変動しているが、Exp1分布およびGP1分布は滑らかに変化している。また、Exp分布およびGP分布のクォンタイルは $r>68$ で乖離しはじめ収束しない。したがって、分布モデル間のクォンタイルの安定性という面では $r=68$ が最適値と考えられる。図中には $r=N_y$ と $r=1.65N_y$ の2本の破線を参考のため示した。1.65 N_y はPOT資料がAMS資料より精度よくクォンタイルを推定できる標本サイズである(Stedinger et al., 1993)。この場合は1.65 N_y は69である。 r の変化に対する確率水文量の変化は2母数と3母数の分布モデルで傾向が異なるので、この傾向が分かれる付近が最適な閾値選定の基準になりうる。さらに、図中にはAMS資料により求めたGEV分布の1/100のクォンタイルも参考までにあわせて示しているが、最適な閾値の判断にAMS資料の確率水文量を基準にすることを条件とするものではない。この理由については第6章において述べる。

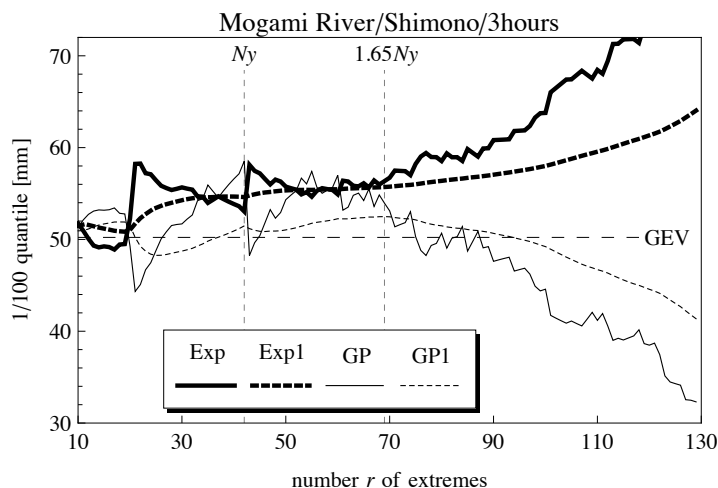


図5.17 1/100クォンタイルの r に対する変化
AMS資料のGEV分布の1/100クォンタイルを破線で示している

2)分布モデルの形状母数および尺度母数

分布モデルの母数による判断は、 r に対する安定性で評価するが、AMS資料の形状母数との一致性を確認しておく。この場合も2母数と3母数の特性の変化が始まる点が重要である。図5.18および図5.19はそれぞれ形状母数と尺度母数の変化を示したものである。両方の図において $r=60$ 付近に比較的安定な領域が存在する。Exp1分布、GP分布、GP1分布の形状母数、尺度母数は $r>68$ において大きく

なる傾向があるが、Exp1分布は r の広い範囲で安定している。したがって、 $r=68$ がこの指標においても最適値と考えられる。POT資料に対するGP分布の形状母数 k はAMS資料に対するGEV分布の形状母数と理論的には互換的であることをふまえ、図5.18にはGEV分布の形状母数も参考までにあわせて示している。なお、後で述べるようにPOT資料の k がAMS資料の k に一致することが適切な閾値選定の条件ではない。形状母数の互換性とクォンタイルの関係については第6章に詳しく述べる。

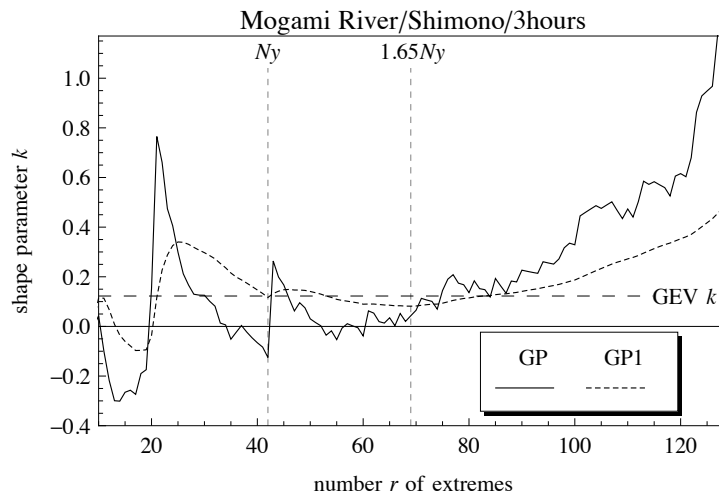


図5.18 形状母数の r に対する変化

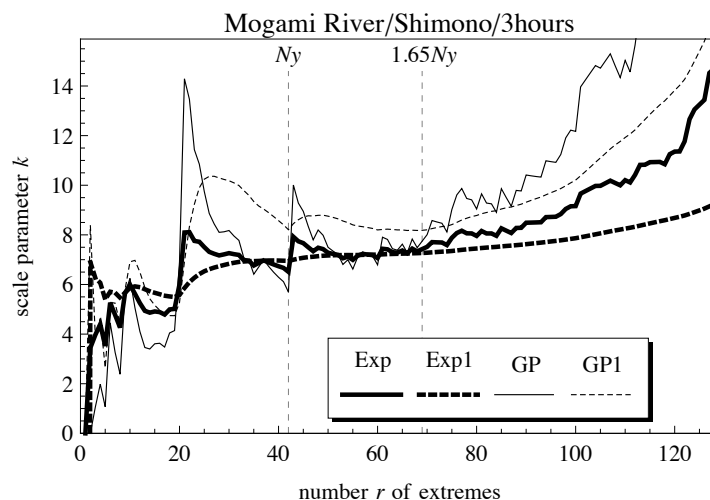


図5.19 尺度母数の r に対する変化

3)SLSC

SLSCは分布の適合度の良否を判断する規準として使われており、地点雨量のAMS資料を対象に他の適合度指標等もあわせて検討された結果、SLSC=0.03が満足すべき適合度の基準とされている（宝・高棹，1988）。SLSCを用いる判断はこの基準値を満たす範囲となる。図5.20に示すように $r=69$ 付近で、SLSC<0.03を満たすのはExp分布、GP分布では $r=74$ が上限である。なお、Exp1分布およびGP1分布についてはかなり広い範囲でSLSC<0.03を満たしている。したがって、SLSCからみると $r=74$ がとりうる r の最大値として考えられる。SLSCは選定すべきでない範囲はわかるが閾値選定の鋭い指標にはならないことが分かる。なお、Exp分布のSLSCが大きくなり始める点 $r=68$ がクォンタイルや母数による判断と一致していることは注目すべきである。

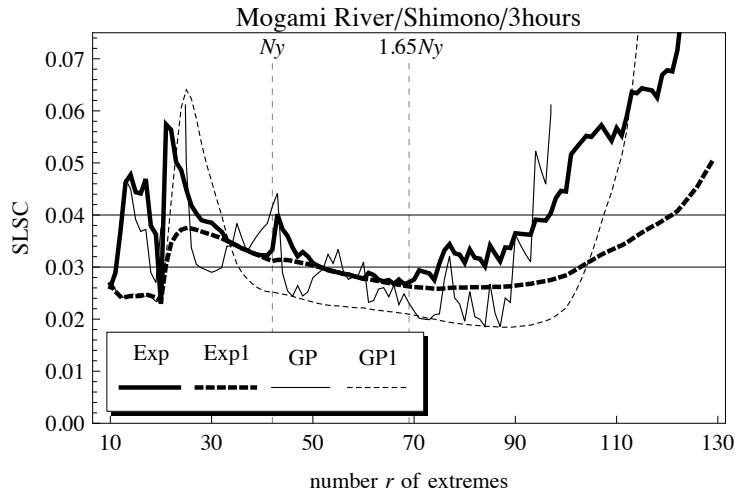


図5.20 SLSCの r に対する変化

4) jackknife推定誤差

jackknife法はリサンプリング手法の一つであり(Efron, 1982), クォンタイル等の推定誤差の変動性を定量化するのに用いられる(宝・高棹, 1988; Takara & Stedinger, 1994). この判断基準は, 値の安定性である. 図5.21に示すように, Exp分布をもとに判断すると $44 < r < 122$ が推奨すべき範囲と考えられるが, 他の分布モデル・母数推定法では閾値選定に関する有用な情報は得られない. また, 一般的にjackknife推定誤差の判断基準を設定することが難しい.

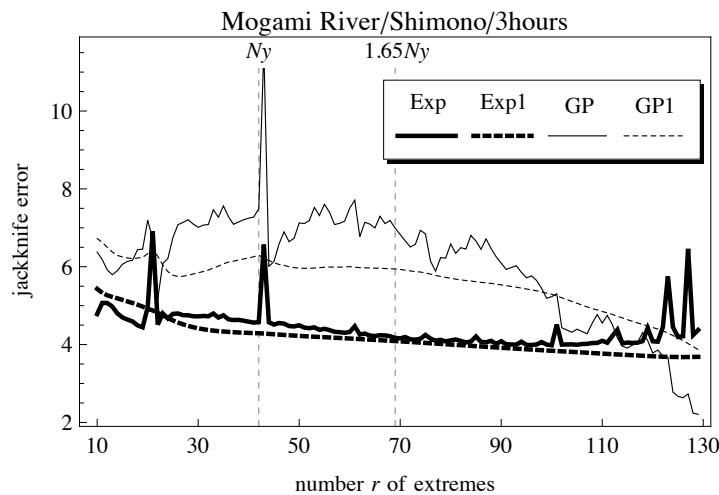


図5.21 jackknife推定誤差の r に対する変化

5) 形状母数を用いた自動選択指標

Reiss & Thomas(1997)は標本サイズ n の標本のうち上位 r 個のデータに対して求められる形状母数 $k_{r,n}$ を用いて自動的に最適な r を選定する方法を提案している. 最適値は次式を最小にする r である.

$$\frac{1}{r} \sum_{i=1}^r i^{\beta} |k_{i,n} - \text{med}(k_{1,n}, \dots, k_{r,n})| \quad \text{with } 0 \leq \beta < 1/2 \quad (5.9)$$

ここに $\text{med}(k_{1,n}, \dots, k_{r,n})$ は $k_{1,n}, \dots, k_{r,n}$ のメディアンである^{注14}.

図5.22および図5.23はそれぞれGP分布およびGP1分布について3種類の β に対して式(5.9)の値の変化を示したものである。図中の縦線は $r > N_y$ の範囲で式(5.9)を最小にする r を示している。これらの図から β が大きい方が閾値選定に際し鋭い指標となることが分かる。GP分布の場合 $r=74$ であり、GP1分布では $r=97$ である。さらに、図5.23より図5.22の方が鋭い判断を与えているので、この自動選択指標については形状母数の算定に用いる母数推定法として上位 r 個の最小値を位置母数とする方法(GP分布)が推奨される。

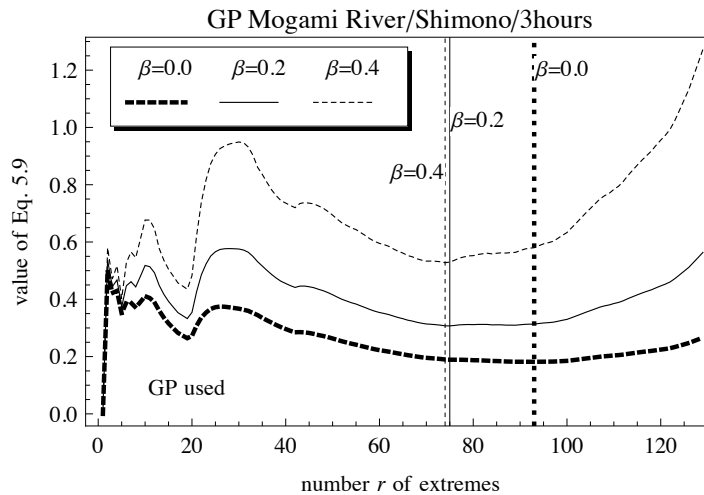


図5.22 GP分布を用いた場合の式(5.9)の値の r に対する変化

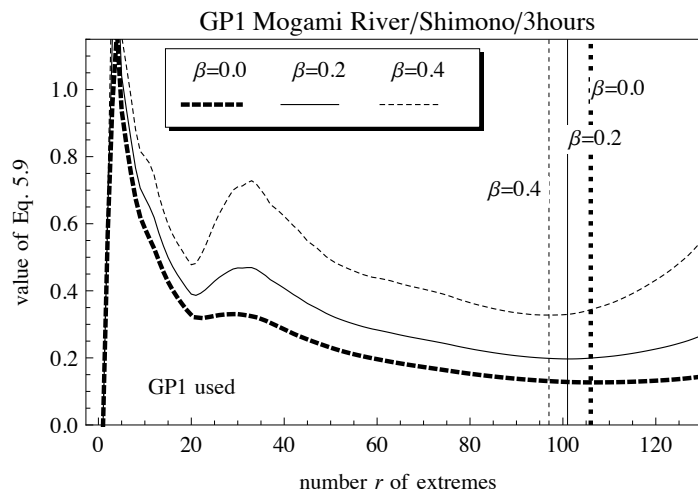


図5.23 GP1分布を用いた場合の式(5.9)の値の r に対する変化

^{注14} Reiss & Thomas(1997)の文献では、形状母数は γ ，上位 r 個は k 個と表示されている。本論文では形状母数に k を用いているので、錯綜を避けるため表記を変えている。

6) 標本平均超過関数(Sample mean excess function, SMEF)

SMEFは次式で定義される関数である(Reiss & Thomas, 1997).

$$e_n(u) = \frac{\sum_{i=1}^n (x_i - u) I(u < x_i)}{\sum_{i=1}^n I(u < x_i)} \quad (5.10)$$

ここにindicator function $I(u < x)$ は $u < x$ のとき1で、それ以外の場合0をとる関数である。この指標による判断は、この式の値が一定となる上限値である。

図5.24はSMEFの変化を示したものである。SMEFは図5.19に示したExp分布の尺度母数によく似ていることが分かる。L積率法によるExp分布の尺度母数は、標本の平均値から標本の最小値を引いて求められる。したがって、式(5.10)において $I(u < x)$ の代わりに $I(u \leq x)$ を用いて、かつ右辺の値から標本の最小値を引いたものに相当する。Reiss & ThomasはPOT資料の閾値（上位 r 個）としてSMEFがほぼ一定となる範囲の上限をとるよう提案しているが、この図からSMEFがほぼ一定となる範囲の上限を判断することは少し難しい。ここでは r として58または68が考えられる。他の標本についてもSMEFはいくつかほぼ一定となる領域が存在し、SMEFのみで最適な上位 r 個を決定するのは難しい。

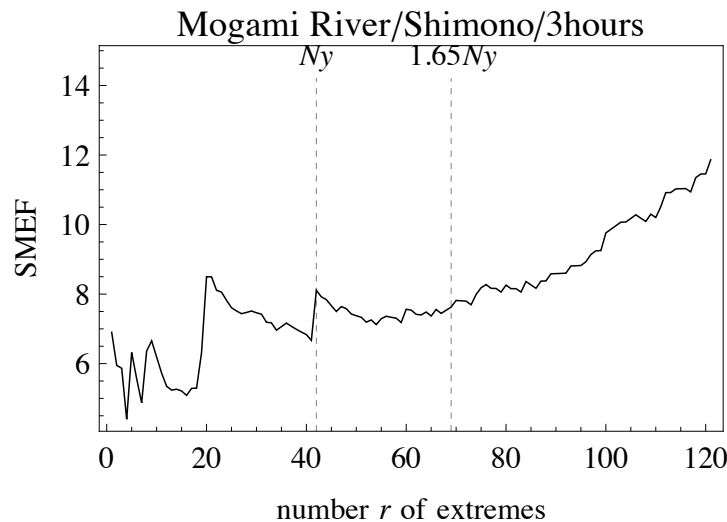


図5.24 SMEFの r に対する変化

7) 指標の比較

上に紹介した6つの指標による推薦する r について表5.1に比較して示した。同表にはあわせて他の標本についての結果も示している。ここでの検討範囲ではすべての標本において、分布モデルの母数が良好な判断を提供しており、閾値選定の有力な指標であると言える。クォンタイルや式(5.9)およびSMEFは少し劣る。結局、下野地点の流域平均3時間雨量の最適な r は68が推奨される。

図5.25および図5.26は下野地点の流域平均3時間雨量のそれぞれ全POT資料($r=129$)および上位68個のPOT資料をAMS資料と比較して示したものである。AMS資料と比較するため、Gumbel確率紙を用い、POT資料は式(3.34)を用いてAMS資料と比較できるように変換してプロットしている。また、AMS資料にはGEV分布を当てはめ、POT資料には上述した母数推定法が異なるGP分布とGPI分布をあてはめ、式(3.29)の変換をおこなってプロットしている。図5.25をみると、全資料を用いた場合、POT資料における15mm以下のデータが極端に曲がってプロットされている。この部分のデータ数が多いため、GP分布およびGPI分布の当てはめは上位のデータと乖離し、これら分布モデルの1/100の

表5.1 閾値選定指標の比較

地点	下野 (3,534km ²)				両羽橋 (6,519km ²)			
降雨継続時間(hrs)	3	6	12	24	3	6	12	24
1/100 Quantile	68	81	70	110	92	104	90	83
Parameters	68	80	62	109	86	104	90	83
SLSC	74	91	100	110	85	104	90	82
jackknife error	44-122	50-120	40-62	70-120	50-110	60-127	90-126	83-119
Equation 5.9 (GP)	74	81	63	110	43	105	92	83
Equation 5.9 (GP1)	97	124	111	122	45	125	113	122
SMEF(main)	58	52	61	74	85	104	89	83
SMEF(sub)	68	80	96	110				111
Recommendation	68	80	62	110	85	104	90	83

クォンタイルは非常に小さくなっており、適当ではないことが分かる。一方、上位68個のデータのみを用いた図5.26ではGP分布およびGP1分布はすべてのデータに良く適合し、適切なクォンタイルを与えていることがわかる。

一方、図5.26で20mm以上の範囲ではAMS資料とPOT資料の間には大きな違いはみられないので、あてはめた分布曲線の傾向が異なるのはAMS資料においてPOT資料の最小値より小さい下位データが影響しているものと考えられる。AMS資料の下位2個の15mm以下のデータがあてはめた分布曲線からはずれていることに注目し、これらがそれより上位のデータの並びと同じようだったらという仮定をおいた時のデータの比較プロットとそれぞれにGEV分布をあてはめたものの比較が図5.27である。下位2個のデータの影響が非超過確率の大きい部分に少なからず影響を与えているのが分かる。図5.25から分かるようにこのレベルの記録はPOT資料にはたくさん見られるので、たまたまAMS資料の下位2個のデータが生じた年の年最大値がこれらのデータであったわけで、ここで仮定したような値をとり得ることも十分あり得るので無理な仮定ではない。田中・宝(2001)はAMS資料の下位データに着目した検討を行っている。これについては第6章で詳しく述べる。

AMS資料では豪雨と言えないような小さな年最大値でも観測期間内にあれば無視することはできないが、POT資料では適切に閾値を選定すれば一つの母集団から抽出された標本とみなせうる大きな極値事象のみを対象に解析することができるという利点を有している。したがって、POT資料による解析はAMS資料に比べ優位であり、もし十分な資料の蓄積があつてPOT資料による解析が可能であれば、適用すべきである。

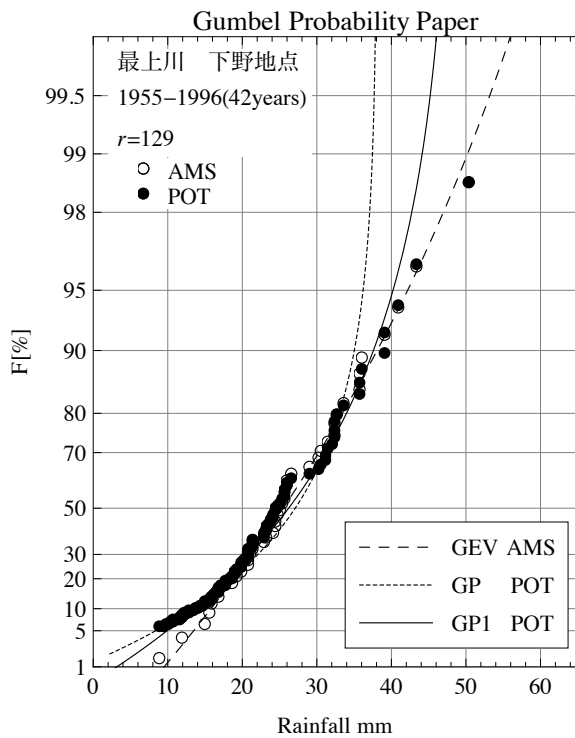


図5.25 最上川下野地点上流流域平均3時間雨量の全POT資料($r=129$)とAMS資料にそれぞれGP分布, GP1分布およびGEV分布をあてはめた結果

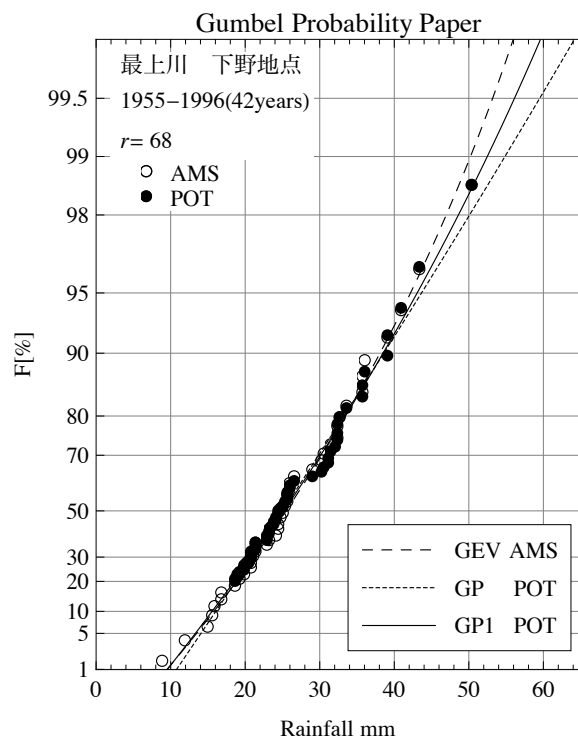


図5.26 最上川下野地点上流流域平均3時間雨量の上位68個のPOT資料とAMS資料にそれぞれGP分布, GP1分布およびGEV分布をあてはめた結果

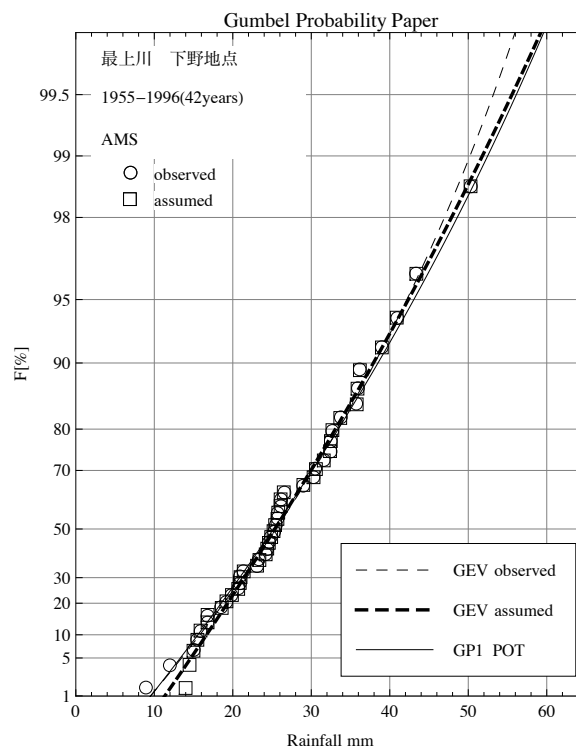


図5.27 最上川下野地点上流流域平均3時間雨量の原AMS資料とその下位2個のデータが少し大きい場合のGEV分布のあてはめの比較

5.5 結論

全国の一級水系の基準地点上流の流域面積雨量の解析から、次のことが言える。

- (1) POT資料を用いる場合閾値の選定により標本サイズが異なってくるが、標本サイズを大きくすることにとらわれすぎるのではなく、確率紙にプロットするなどにより上位データと同じような特性を有する範囲を選ぶ必要がある。
- (2) POT資料はAMS資料に比べ正に偏る形状母数を与える傾向にある。この傾向はPOT資料の標本サイズがAMS資料より大きい場合より顕著である。
- (3) 形状母数の地域内平均値は北から南へ向かうにしたがって大きくなり正に偏る傾向がある。また、どの地域においてもPOT資料の形状母数がAMS資料の形状母数より大きい結果となっている。
- (4) L積率比でみるとPOT資料は指数分布および一般パレート分布があてはまるものが多く、AMS資料はGumbel分布およびGEV分布があてはまるものが多い。

全国の一級水系の79地点の洪水ピーク流量資料について、適合度と安定性から評価するとともにAMS資料とPOT資料の両面から評価することにより以下のことがわかった。

- (1) AMS資料と同じ標本サイズとなるように抽出したPOT資料（非毎年年数最大値）は満足すべき適合度の基準として $SLSC \approx 0.04$ をとると適合度を満足しなくなる標本が少なからずある。
- (2) POT資料で評価することにより、洪水規模の均質性が確保されるとともに、水位流量曲線の接続点周辺のデータのバラツキ等の問題を回避できる場合がある。
- (3) 適合度評価、リサンプリング手法と併せてAMS資料とPOT資料を併せて評価することにより、確率水文量の信頼性を向上させることができる。
- (4) 観測期間が連続しない資料について、POT資料による頻度解析が有力な手段となり得る。

流域平均雨量を対象にした閾値の選定方法に関する検討では、次のことがわかった。

- (1) 確率水文量の変化傾向が分かりやすい指標となりうるが確率水文量単独で r を特定するのは難しい。分布モデルの母数による判断指標が良好な判断を与える。
- (2) 選定された閾値を用いて抽出したPOT資料にあてはめた分布モデルは非超過確率の大きい範囲のデータに対しても良好に適合する。一方、AMS資料の下位データが非超過確率の大きい範囲のクォンタイルに影響を及ぼしている場合がある。
- (3) わが国の水文資料を対象にする場合、POT資料による解析はAMS資料による解析に比べ優位であるので、従来のAMS資料を用いた解析に加え、POT資料による解析を行うべきである。

参考文献

- Efron, B. : The Jackknife, the bootstrap and Other Resampling Plans, SIAM Monograph, No.38, p.92, 1982.
- Rasmussen, P.F., F. Ashkar, D. Rosbjerg, and B. Bobee: The POT method for flood estimation: a review, Stochastic and Statistical Methods in Hydrology and Environmental Engineering, Vol.1, pp.15-26, 1994.
- Reiss, R.-D. and M. Thomas: Statistical Analysis of Extreme Values, Birkhäuser, p.316, 1997.
- Stedinger, J.R., R.M. Vogel, and E. Foufoula-Georgiou: Frequency Analysis of Extreme Events, Chap. 18, Handbook of Hydrology, (Ed.) D. R. Maidment, McGraw-Hill, New York, pp. 18.1-18.66, 1993.
- Takara, K. and J.R. Stedinger: Recent Japanese contribution to frequency analysis and quantile lower bound estimator, Extreme Values: Floods and Droughts(ed. by K. W. Hipel) Stochastic and Statistical Methods in Hydrology and Environmental Engineering, Kluwer, Dordrecht, The Netherlands, vol.1, pp.217-234, 1994.
- Tanaka, S. and Takara, K.: A Study on Threshold Selection in POT Analysis of Extreme Floods, The Extremes of the Extremes, IAHS Pub. no.271, pp.299-304, 2000.
- 江藤剛治, 室田 明, 米谷恒春, 木下武雄: 大雨の頻度, 土木学会論文集, 第369号/II-5, pp. 165-174, 1986.
- 角屋 睦: 水文統計論, 水工学シリーズ, 土木学会水理委員会, 64-02, p.59, 1964.
- 宝 馨, 高埴琢馬: 水文頻度解析における確率分布モデルの評価規準, 土木学会論文集, 第393号/II-9, pp.151-160, 1988.
- 田中茂信, 宝 馨: 毎年資料と非毎年資料による確率水文量の評価, 水工学論文集, 第43巻, 土木学会水理委員会, pp.145-150, 1999a.
- 田中茂信, 宝 馨: 河川流量の頻度解析における適合度と安定性の評価, 水工学論文集, 第43巻, 土木学会水理委員会, pp.127-132, 1999b.
- 田中茂信, 宝 馨: 洪水頻度解析におけるAMSとPDSの比較, 水工学論文集, 第45巻, 土木学会水理委員会, pp.205-210, 2001.

6. AMS資料とPOT資料の比較

前章までにAMS資料およびPOT資料それぞれを用いた頻度解析について述べてきた。AMS資料とPOT資料のいずれを用いるべきかという問題は古くから議論されているが、それぞれに一長一短があり結論が出ていない。また、それぞれの方法を単独に使用する限りあまり問題は見えてこない。一つの時系列資料から得られたAMS資料とPOT資料からは同じ確率水文量が期待されており、理論的にはほぼ同じ確率水文量を与えるものである。わが国では、主に資料整理が簡便なAMS資料の解析が慣用されてきており、多くの研究や応用が実施されてきた実績がある（たとえば、水文・水資源ハンドブック、1997）が、POT資料を用いた解析については角屋(1964)などに方法論が紹介されているが、実務的にはこれまで十分な知見が得られていない。第5章の終わりに示したようにPOT資料を変換してAMS資料と比較すると、AMS資料による解析にも問題があることが鮮明になる。本章ではPOT資料にGP分布をあてはめAMS資料にGEV分布をあてはめるとき、両者の間で形状母数 k が互換的に使用できるという特性に基づいてわが国の水文資料の特性をみてる。そして、AMS資料とPOT資料の間で形状母数が互換的に使用できるものばかりでない、というより、互換的に使用できるものの方が少ないことを示すとともにその原因を探り、解決方法を提案する。

6.1 AMS資料とPOT資料による同一時系列資料の解析

同一時系列資料から得られたAMS資料とPOT資料を用いた解析では、最大極値に対して同じ確率水文量が得られると期待される。両者における事象の頻度の間には、第4章、第5章で述べたようにいくつかの一般的な関係が存在する。ここでは第5章5.1で述べた理由により、POT資料にGP分布および指数分布(Exp分布)をあてはめAMS資料にGEV分布およびGumbel分布をあてはめて両者を比較しながら解析する。

6.1.1 解析に使用した資料

ここでは、わが国の一級河川における105地点の洪水ピーク流量の時系列資料から得られたAMS資料とPOT資料を用いた。観測期間の平均値は42年である。最大は67年で最小は27年である。POT資料については、第5章5.4で述べたようにいろいろな観点から検討して閾値設定を行う必要があるが、ここでは非毎年年数最大値（POT資料のうち統計年数と同じ数だけ上位から取り出したもの）を使用する。

6.1.2 AMS解析とPOT解析による確率水文量

解析では分布モデルの母数推定はすべてL積率法により行った。また、AMS解析の確率水文量と比較するPOT解析の確率水文量を算出する際には、式(4.1)の関係からAMS資料の年超過確率に相当するPOT資料の超過確率を求め、それを使用している。

図6.1は年超過確率1/100の確率水文量をAMS資料にあてはめたGEV分布とPOT資料にあてはめたGP分布について比較したものである。GP分布はGEV分布より小さめの確率水文量を与えることが多い。GEV分布の確率水文量に対するGP分布の確率水文量の比について平均をとると0.925となる。

図6.2は図6.1と同様に年超過確率1/100の確率水文量をAMS資料のGumbel分布とPOT資料のExp分布について比較したものである。Exp分布はGumbel分布より大きめの確率水文量を与えることが多く、Gumbel分布の確率水文量に対するExp分布の確率水文量の比について平均をとると1.062となる。

GEV分布とGP分布は3母数の確率分布である。一方、Gumbel分布とExp分布は2母数の確率分布である。図6.1と図6.2から、2母数と3母数の分布ではAMS資料とPOT資料で確率水文量の大小関係が異なることが分かる。この理由については後ほど確率紙による検討で述べる。

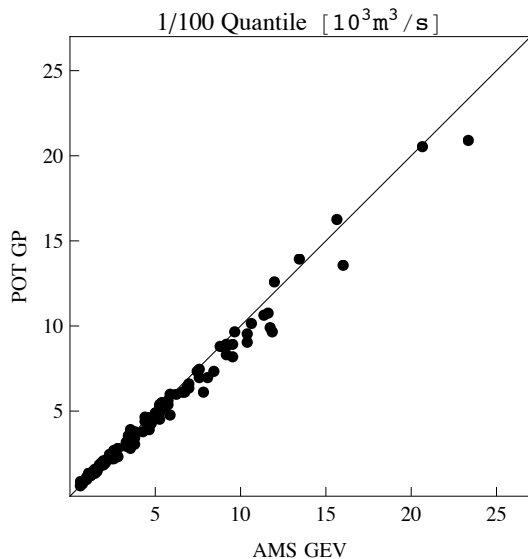


図6.1 AMS(GEV)とPOT(GP)の確率水文量の関係

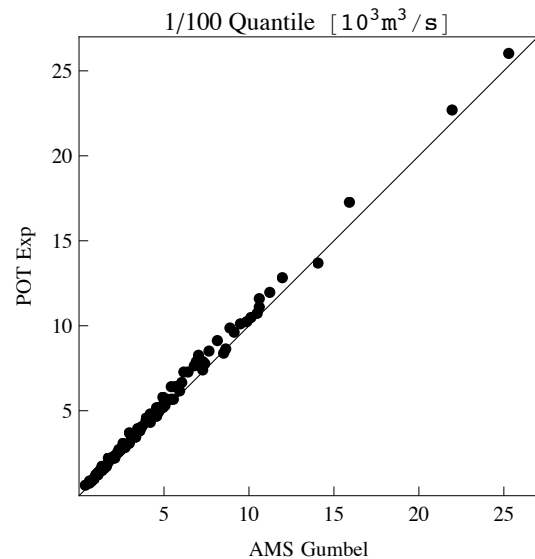


図6.2 AMS(Gumbel)とPOT(Exp)の確率水文量の関係

6.1.3 AMS解析とPOT解析による母数

1) AMS解析とPOT解析によるそれぞれの形状母数および尺度母数の比較

前述したように、理論的には形状母数はAMS資料とPOT資料で互換性を有している。図6.1および図6.2で示した確率水文量の傾向が異なる原因を調べるためAMS資料およびPOT資料の形状母数の特性について調べてみる。

図6.3はGEV分布とGP分布の形状母数 k を比較したものである。全資料（105標本）のうち $k > -1/3$ （ k がひずみ係数と理論的關係を有する範囲）の条件を満たしかつAMS資料とPOT資料の間で k の値の差が0.05未満となるものを●印で、それ以外を○印で示している。●印は17標本みられるが、他の大部分は k がAMS資料とPOT資料でかなり異なる。GEV分布の k は0.2より小さい部分に分布しており、平均値は-0.129であるのに対し、GP分布の k は正負にバランスよく分布しており、平均値は0.026である。したがって、一般的にGumbel確率紙上でGEV分布は上に凸の曲線となり、裾の長い分布となるのに対し、GP分布はほぼ直線であるがやや下に凸の曲線となり、上限を有する分布となることが多い。

図6.4は同様の比較をGumbel分布とExp分布の尺度母数について行ったものである。●印は図6.3と同じ条件で形状母数を規準に選んだ17標本を対象にしている。ほとんどの資料がAMS資料およびPOT資料で同じ尺度母数を与えるかAMS資料のGumbel分布の方が少し小さい尺度母数となっている。また、形状母数が互換性を有している標本は尺度母数でみても互換性を有している。尺度母数が小さいと同じ非超過確率に対し、小さい確率水文量を与えることになる。

このようにAMS資料およびPOT資料を用いて推定した母数も6.1.2の結果を裏付けている。

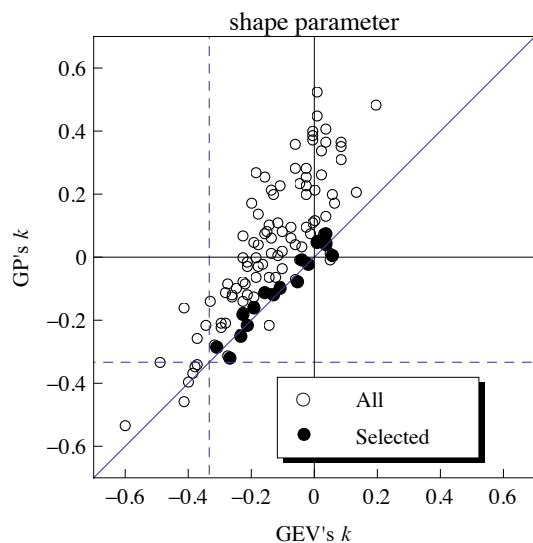


図6.3 AMS(GEV)とPOT(GP)の形状母数 k の比較

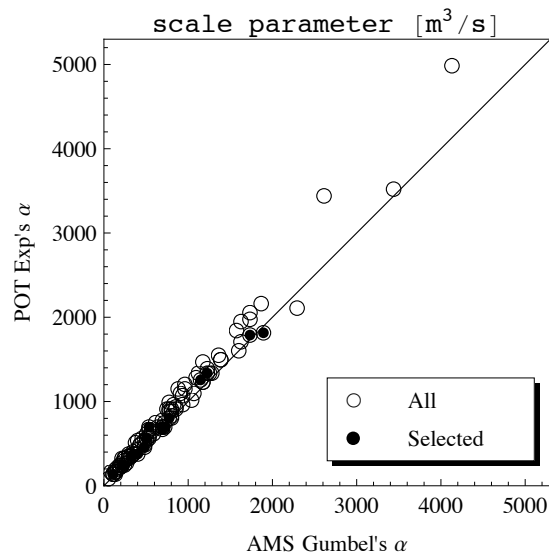


図6.4 AMS(Gumbel)とPOT(Exp)の尺度母数 α の比較

2)形状母数がAMS資料とPOT資料で互換性を有する標本の特性

図6.1の確率水文量の図において形状母数 k が上述の条件で一致しているとみられる17標本を●印で区別して示したものが図6.5である。●印についてはGEV分布とGP分布の確率水文量がよく一致している。すなわち、 k が一致すると確率水文量も一致することが見てとれる。

次に、これらの17標本に対するGEV分布とGP分布の適合度をSLSC（宝・高棹，1988）により評価する。図6.6はGEV分布のSLSCとGP分布のそれを比較したものであり、全資料との比較で示している。選択された17標本のSLSCは河川流量の適合度の基準である0.04（田中・宝，1999）以内である。

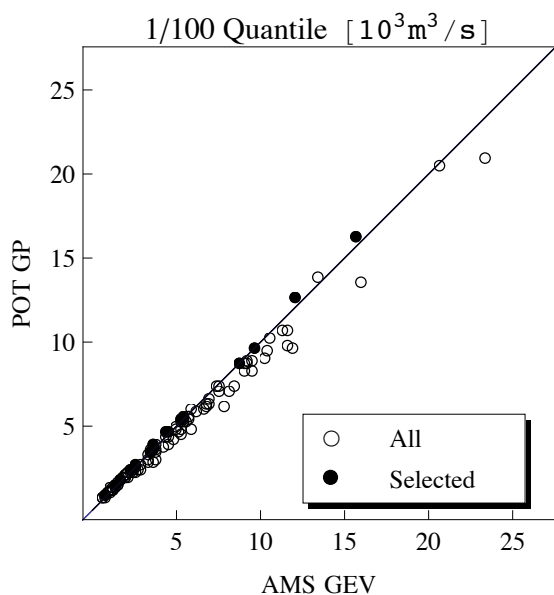


図6.5 形状母数 k が一致する標本に関するAMS(GEV)とPOT(GP)の確率水文量の関係の全資料との比較

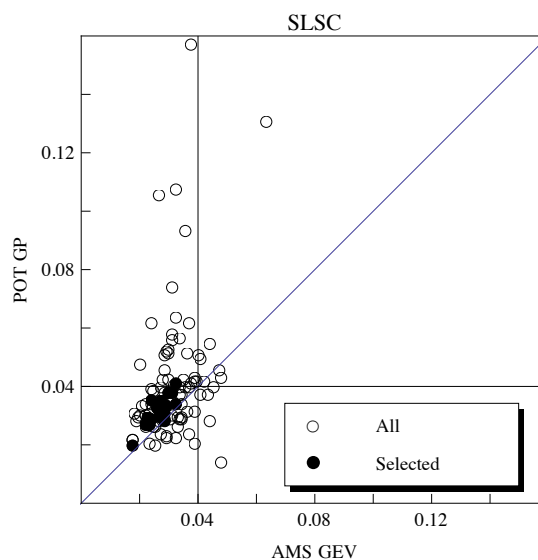


図6.6 形状母数 k が一致する標本に関するAMS(GEV)とPOT(GP)のSLSCの関係の全資料との比較

確率分布モデルの安定性の評価はリサンプリング手法(Efron, 1982)で行う。リサンプリング手法にはjackknife法やbootstrap法があるが、ここでは jackknife法を用いる。jackknife法による確率水文量の推定誤差を不偏推定値で除した相対誤差を確率水文量の安定性評価の指標として用いることができる。図6.7は相対誤差をGEV分布とGP分布で比較したものであり、全資料とともに示している。選択された17標本のうち16標本がAMS資料およびPOT資料ともに相対誤差0.25未満であり、全体の中でも安定性が良好である。

以上の結果から、形状母数 k がAMS資料とPOT資料で一致する標本の場合、確率水文量がAMS資料およびPOT資料いずれで評価しても同程度となり、その上、適合度や安定性もAMS資料およびPOT資料いずれの場合においても良好である。すなわち、このような場合は、AMSとPOT資料のいずれを用いても同様の結果が得られ、どちらを用いてもよい。

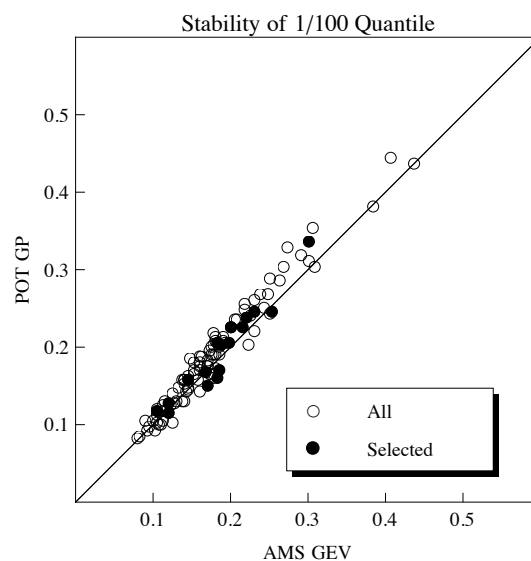


図6.7 形状母数 k が一致する標本に関するAMS(GEV)とPOT(GP)のjackknife相対誤差の関係の全資料との比較

3)形状母数 k が互換性を有していない標本の特性

形状母数 k がAMS資料とPOT資料で互換性を有していない時系列資料が84%と大部分を占めており、これらについてAMS資料とPOT資料のいずれを選択すべきかについて検討する必要がある。確率水文量、適合度、安定性については図6.4～図6.6で示したように k が互換性を有する標本の特性と対比される。

ここでは、 k が互換性を有していない代表的な標本について確率紙上でどのようにプロットされるのかをみている。

図6.8はAMS資料とPOT資料の標本をGumbel確率紙にプロットしたものである（プロットングポジションはカナンプロットを用いている）。この標本に対するGEV分布の $k=-0.209$ であり、破線のような上に凸の形状となる。GP分布の $k=-0.017$ であるのでほぼ直線（若干上に凸）である。GEV分布のSLSC=0.029で、GP分布のSLSC=0.037である。ここではPOT資料の標本とExp分布およびGP分布は式(4.1)の関係を用いてAMS資料と比較できるように変換してプロットしてある。AMS資料、POT資料ともに大きい事象はほぼ同じ位置にプロットされる。しかし、小さい事象（小洪水）の部分でAMS資料とPOT資料は図のように異なる。第5章5.4.2の図5.27で示したように、AMS資料の下位のデータ付近でPOT資料と傾向が異なることが確率分布のあてはめに大きな影響を与える。多くの

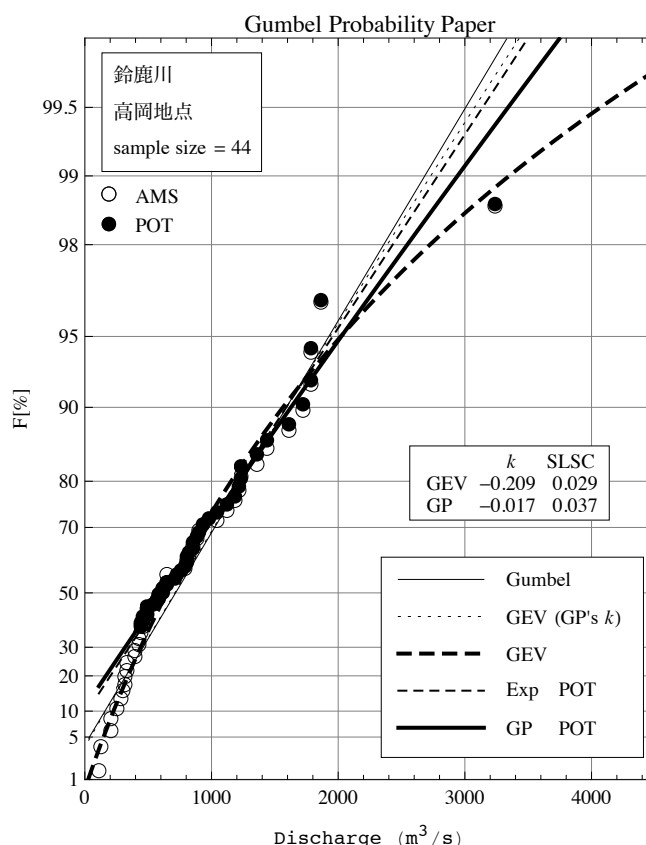


図6.8 AMS(GEV)とPOT(GP)の形状母数 k が一致しない
標本のプロットと確率分布モデルの比較

AMS標本ではこの図のように小さい洪水ピーク流量がほぼ同じ値となり、確率紙上で縦に並ぶ傾向があるが、POT標本ではこのようになることは少ない。この影響でAMS資料とPOT資料の分布のあてはめが以下になることが多い。まず、2母数の場合、AMS資料にあてはめたGumbel分布はほぼ縦に並んだ下位データの影響でPOT資料のExp分布に比べ少し反時計回りに回転したグラフとなることが多く、結果としてExp分布に比べて小さめの確率水文量を与える。一方、3母数の場合、AMS資料にあてはめたGEV分布は上に凸形となるのに対し、POT資料のGP分布はほぼ直線または少し下に凸形になることが多く、形状母数 k は前者の方が一般に小さくなり、非超過確率の大きな事象（洪水）に対してはより大きな乖離がみられるとともに概してGEV分布の確率水文量はGP分布の確率水文量より大きく算定されることになる。

このように、確率紙へのプロットは適合度や安定性を示す数値以上に説得性を有しており、多くの情報が得られるので、単にモデルをあてはめて確率水文量を計算するのではなく、AMS資料とPOT資料を確率紙にプロットして総合的な判断をすることも重要である。

4) POT資料の優位性

ここでの検討はPOT資料については非毎年年数最大値としているが、他の問題が起きない範囲で取り扱う極値データの個数を増やすように閾値を設定することにより信頼性の高い推定を行うことが可能である。また、中小洪水の影響でどうしても適合度が良くない場合（たとえば、洪水流量の観測値は規模により精度や傾向が異なる場合がある）には、それらより大きな洪水のみを対象とするように閾値を設定することができるなどの柔軟性も有している。

6.1.4 形状母数の互換的使用による確率水文量

多くの一級河川では昭和30年前後から観測が継続的に行われているが、わが国で現在一般に求めようとする確率水文量の再現期間は100～200年であり、どうしてもパラメトリックな確率モデルにより外挿的に確率水文量を推定することになる。上述したように、AMS資料は、年最大値といえども渇水年や台風の少ない年には洪水と呼べないような値である場合があり、また、出水の多い年の年間第2位の出水で他の年の最大値より大きいものが解析の中に取り込まれないという問題があることが指摘されている。一方、POT資料では、資料年数の短さとAMS資料の問題点を解消することができる可能性がある。ここではPOT資料を用いて得たGP分布の k をAMS資料の解析に使用できるかどうか検討する。

図6.9は、横軸に通常のGP分布、縦軸はPOT資料に対するGP分布の k を使用して他の母数をAMS資料に対して推定したGEV分布それぞれの確率水文量の比較である。同じ k を使用してほぼ同じ確率水文量を与えており、GP分布の k を互換的に使用することができる。また、図6.1に比べて確率水文量のばらつきが少なくなっている。

図6.10はGEV分布の確率水文量を k の違いで比較したものである。横軸は通常のGEV分布であり、縦軸はPOT資料に対するGP分布の k を使用し、他の母数をAMS資料に対して推定したGEV分布である。通常のGEV分布は後者に比べ少し大きめ（105標本の平均で1/100確率水文量が1.15倍）の確率水文量を与える。図6.9、図6.10から、 k が確率水文量の推定に際して大きな役割を果たしていることが分かったとともに適切な k を設定することが重要であることがわかる。

6.1.5 AMS資料とPOT資料における形状母数の一般的傾向と対策

図5.9に示したように、非毎年年数最大値とAMS資料の上位データは既往最大値を含め最大約4割は一致している。したがってPOT資料を非毎年年数最大値とする限り（AMS資料とは異なる観測期間の場合、最大値が異なるだけでなく、プロットされる位置も変わってくるので比較が難しい）、上位データの先にある外挿範囲での超過確率やクォンタイルはほぼ一致することが期待される。形状母数はAMS資料に対しGEV分布をあてはめたものとPOT資料にGP分布をあてはめたものが理論的には相互に互換的に利用できるはずであるが、図6.3および図5.2、図5.4、図5.5、図5.12で見てきたように

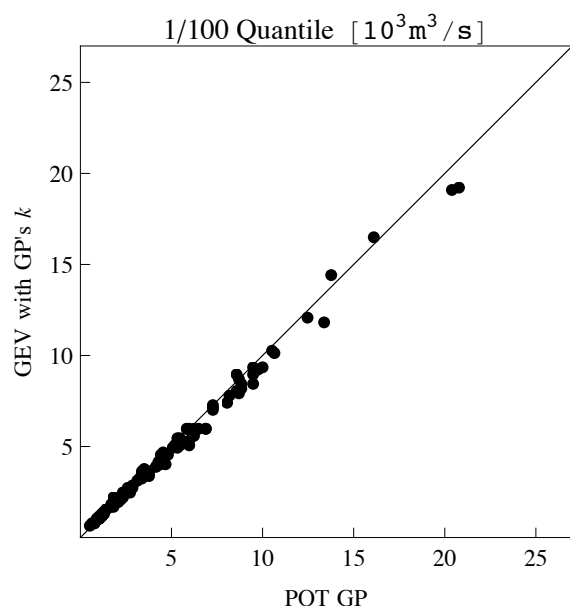


図6.9 GP分布の k を用いたGEV分布とGP分布の確率水文量の比較

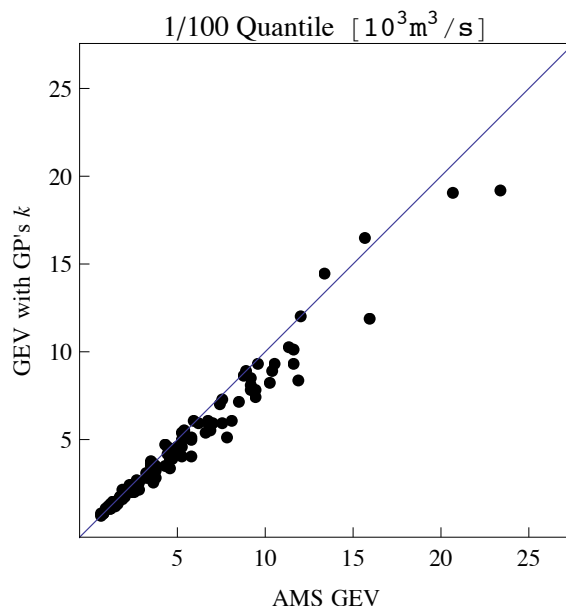


図6.10 AMS資料の同一標本に対しGEV分布とGP分布の k を用いたGEV分布の確率水文量の比較

AMS資料のGEV分布の k よりPOT資料のGP分布の k の方が大きい。 k の違いは外挿範囲でのクォンタイルに少なからず影響を与える。3母数のGEV分布とGP分布では、AMS資料を用いるGEV分布の方が大きめの確率水文量を与え、逆に2母数のGumbel分布と指数分布ではAMS資料を用いるGumbel分布の方が小さい確率水文量を与える。多くの場合、6.1.3に示したことが原因となって、形状母数が異なれば2母数の分布での尺度母数もわずかではあるが異なることになる。このような場合、5.4で示した方法により適切な閾値を選定したPOT資料を用いて確率水文量を求めるべきである。この際、AMS資料を用いた確率水文量や形状母数などは参考程度に参照されるべきものであり、AMS資料の値と近いものを選ぶことにはならないことに注意しなければならない。

6.1.6 まとめ

以上の結果から重要な点を挙げれば(1)~(3)のようになる。

- (1)確率水文量の推定に際して k が重要な役目を果たしており、適切に k を設定する必要がある。
- (2) k がほぼ同じであればAMSのGEV分布とPOT資料のGP分布はほぼ同じ確率水文量を与え、適合度および安定性についても良好である。
- (3)POT資料のGP分布の k よりもAMS資料のGEV分布の k の方が小さい事例が多く、AMS資料の k の分布はPOT資料に比べて負に偏っている。このことは確率紙でも確認できる。すなわち、AMS資料、POT資料ともに超過確率の小さい事象はそれぞれの順序統計においてだいたい共通であるが、AMS資料においては下位データにおいて同程度の小さい洪水ピーク流量がほぼ縦に並び、Gumbel確率紙上でより勾配が大きくプロットされる。そしてこれがPOT資料に比べAMS資料の k を負に偏らせ大きめの確率水文量を与える原因になっている。

従来から指摘されているように、AMS資料の場合、渇水年における年最大値が洪水と呼べないような小さな値になることがある。また、年最大値のみを扱うので年間の第2位の洪水ピーク流量が解析に用いられない。一方、POT資料においては閾値を適切に設定すれば“洪水”のみを対象とした解析ができる。さらに、観測が不連続な場合でも、閾値を適切に設定すれば連続しない資料も含めた全資料を解析に用いることができる。

したがって、わが国の洪水頻度解析においては、当面、AMS資料だけでなくPOT資料による解析も同時に行い、両者の一致性を見た上でどちらを選択するか判断することが望ましい。AMS資料とPOT資料が異なる傾向の時は、閾値が適切に設定されたPOT資料が用いられるべきである。なお、POT資料はAMS資料に比べ多くの利点を有しているが、独立なピークの抽出と閾値の設定に関する課題が残っている。前者についてはわが国の場合、洪水が数日で終結し、数ヶ月も続くことがないのであまり問題にならない。閾値設定のガイドラインを策定することがPOT資料の最大の課題となっており、いろいろな取り組みが行われている(Rasmussen, et.al, 1994, Tanaka &Takara, 2002)。

6.2 形状母数が異なる原因および考えられる対応策

6.2.1 極値の特性と極値分布

そもそもわれわれが取り扱おうとしている極値は極値分布に従うのであろうか。極値とは母集団から独立に抽出された標本の最大値または最小値などのことである。極値の分布は母集団から取り出す標本のサイズ n が大きくなると極値分布に近づくことが知られている。通常、極値分布とはこの極限形式を意味する。極値分布は3つの形式に分類され、それらをまとめて一つの式で表したものが一般極値分布（GEV分布）である。極値が抽出されたもとの母集団が従う分布が、指数分布、ガンマ分布、正規分布、対数正規分布、ロジスティック分布のいずれであってもその極値はGumbel分布に従

う（志村，2008）．AMS資料はたとえば1年で区分されるよう抽出された標本の最大値からなる標本である．したがって理想的なAMS資料の場合はGumbel分布またはそれより柔軟性が高いGEV分布に従うと考えられ，このAMS資料を抽出したもとの資料から得られるPOT資料は理想的には指数分布またはそれより柔軟性が高いGP分布に従うと考えられる．

しかしながら，これまでみてきたように実際にはこの関係を満たす資料はそう多くない．この原因として，一つは年間の洪水のイベント数が少ないことが考えられる．極値理論はもともと1年間などの1区分内のイベント数 n が十分大きい場合に極値分布に近づくのであって，これが少ない場合には極値分布（GEV分布）に十分近づかない可能性がある．これは簡単なモンテカルロシミュレーションによって確かめることができる．

図6.11は母集団が指数分布に従う条件の下で，年間 $n=10$ および360のイベントがあることを想定し，40年，60年，100年分のイベントを発生させ，その資料から年最大値および非毎年年数最大値を抽出したもののについて，それぞれL積率比 τ_3 と τ_4 の関係をプロットしたものである．各条件についてそれぞれ100ケース実施した結果である．図より1年あたりの洪水発生数 n はここでの検討範囲ではあまり顕著な影響は与えていない．AMS資料はGumbel分布に十分近づいていないが，そのばらつきの範囲は概ねGEV分布に沿っている．POT資料も同様に指数分布に十分近づいていないが，そのばらつきの範囲はGP分布に沿うとともにAMS資料ほどは広く分布していない．40年の場合，一般ロジスティック分布(GLO)から一般パレート分布(GP)の範囲の外にも少なからず分布しているが，年数が60,100と増えるにつれてばらつきが少なくなるのが分かる．一方，100年の場合では，AMS資料はGumbel分布，POT資料は指数分布を中心に分布し，かつ，それぞれGEV分布およびGP分布の曲線に沿ってデータがプロットされているのがわかる．40年の場合のデータの分布は，図5.8に示した分布状況と類似していることがわかる．なお，この $\tau_3-\tau_4$ 関係図においてAMS資料のGEV分布のからの誤差のRMSEに比べ，POT資料のGP分布からの誤差のRMSEは0.74～0.86と常に小さくなっている．このことはAMS解析よりPOT解析が有利であることを示唆している．

これと同様の検討を母集団が正規分布にしたがう場合について行ってみたが，図6.12に示すようにAMS資料のGumbel分布への漸近，POT資料の指数分布への漸近は母集団が指数分布の場合より遅い．AMS資料のプロットはGumbel分布より正規分布に近いものが多いのに比べ，POT資料は100年の場合，指数分布にかなり近くなっているとともに，GP分布の曲線にかなり近くなっており，母集団が指数分布の場合とあまり大きな違いがなくなっている．

図6.13は図6.11の資料についてAMS資料に対してGEV分布をあてはめた形状母数 k と非毎年年数最大値にGP分布をあてはめた形状母数 k を求めて比較したものであり，年間のイベント発生回数10回で40年の場合と100年の場合を比較している．40年に比べ100年の方は形状母数のばらつきが小さくなっている．この図では第5章図5.5および5.12で示したようにGP分布の形状母数がGEV分布の形状母数より大きめとなる傾向はあまり明白ではない．図6.13の平均をとると，40年の場合でPOT資料の形状母数がAMS資料の形状母数より0.05ほど大きい．

以上のことから，L積率比については，観測期間が短いことによってばらつくことは説明できそうであるが，形状母数が異なることについては，観測期間の長短の問題ではないようであるので，記録の蓄積が100年程度まで進んでも解決されるものではない．

しかしながら，実際の観測データではAMS資料とPOT資料で形状母数が異なる特性を有している．これについては次節で述べる複数年で最大値を抽出した標本を考えることにより，1区分内でのイベント数ある程度確保するとともにAMS資料における下位資料に同じようなデータが並ぶのを防ぐことができる（田中，2004）．

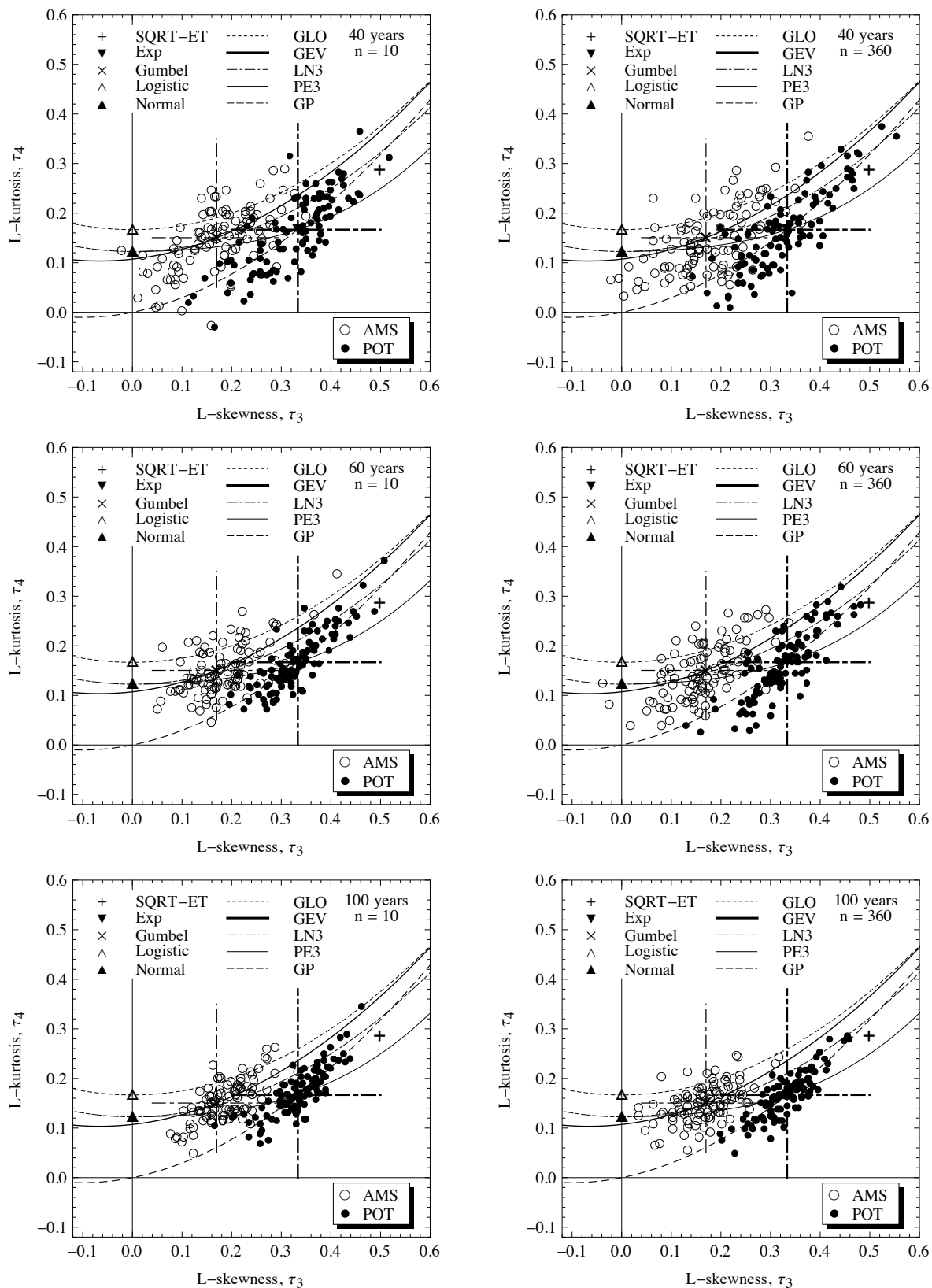


図6.11 モンテカルロシミュレーションによるAMS資料とPOT（非毎年年数最大値）資料のL積率比の比較
（母集団が指数分布の場合）

太い一点鎖線の十字は指数分布，細い一点鎖線の十字はGumbel分布の τ_3 - τ_4 関係を表している。

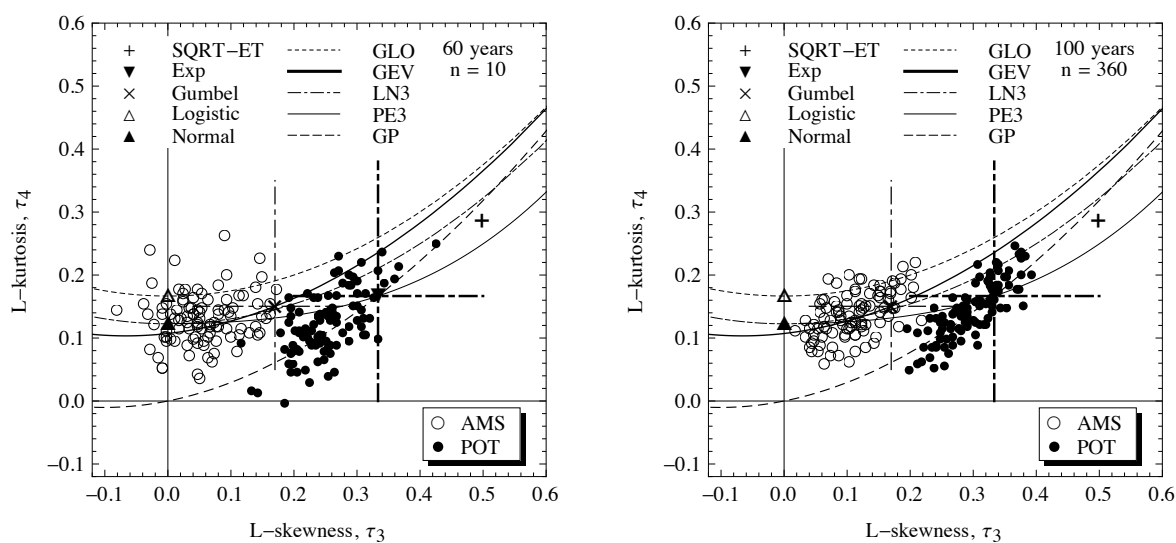


図6.12 モンテカルロシミュレーションによるAMS資料とPOT（非毎年年数最大値）資料のL積率比の比較（母集団が正規分布の場合）

太い一点鎖線の十字は指数分布，細い一点鎖線の十字はGumbel分布の τ_3 - τ_4 関係を表している。

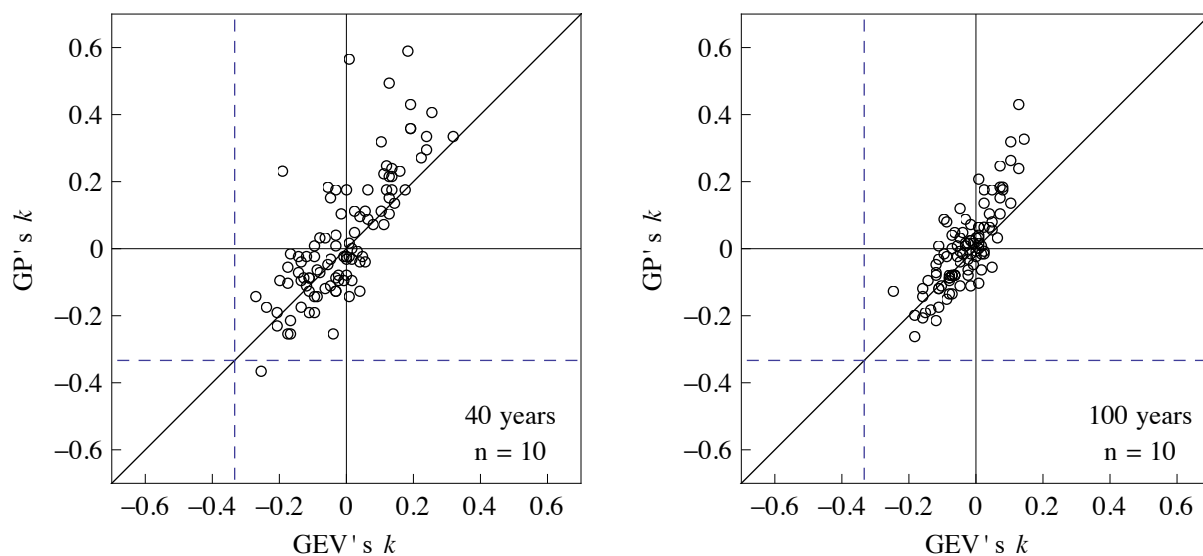


図6.13 モンテカルロシミュレーションによるAMS資料とPOT（非毎年年数最大値）資料の形状母数の比較（母集団が指数分布の場合）

それ以外の原因として，標本の均一性が保たれていないことが考えられる，これはPOT資料で見ると，小さい洪水や降雨がそれより上位のイベントとは異なる傾向を持つ場合などに相当し，この影響を除くにはPOT解析の際に適切に閾値を選定する必要がある，第5章で示した方法により対応できる。

6.3 複数年で区切った最大値資料の解析

AMS資料とPOT資料の比較については、田中・宝(2001)が、POT資料で整理するとAMS資料とは異なる確率水文量を与える場合があること、また、洪水規模の小さいイベントの発生状況によってはPOT資料に比べAMS資料の方が大きめの確率水文量を与える場合があることを指摘している。さらに西岡・寶(2002)は、洪水の発生過程が季節性を持つことを指摘し、AMS資料を用いたGEV分布が過大な確率水文量を与える場合があることを指摘しており、水文事象の1年間の平均生起個数が4個程度以上であるか、AMS解析とPOT解析による確率水文量が一致する場合に、GEV分布は精度の良い確率水文量を与えることを示した。

ここでは、これらの指摘を踏まえ、下位データの影響を除くべく時系列を1年で区切って整理したAMS資料に対し、2年、3年などの複数年で区切って整理し、AMS資料やPOT資料と比較検討し、その特性を評価する。

6.3.1 資料の整理方法

一般に洪水のピーク流量の年最大値には洪水と呼べないようなデータが含まれる。図6.14は豊川の事例であり、この図の範囲では1980年の554m³/sが最小値である。しかしながら、小さいデータはそう何年も続いて出現してはいない。1,000m³/sより小さいデータを対象に考えると、この事例では1980年、1981年と続いている以外はみられず、他は単独で出現している。

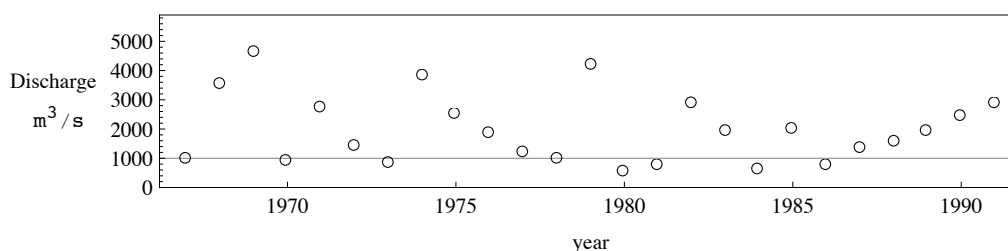


図6.14 豊川の洪水ピーク流量AMS資料の時系列

田中・宝(2001)は洪水ピーク流量のAMS資料にはこのように小さいデータの出現状況がそれより大きな中規模の洪水とは異なることを指摘した。また、このような資料をPOT資料と比べると、AMS資料のGEV分布の方がPOT資料のGP分布より大きな確率水文量を与えることを示した。

(AMS資料における小さいデータの出現状況がそれより大きな中規模のデータと出現状況と異ならない場合は、AMS資料のGEV分布の形状母数とPOT資料のGP分布の形状母数がほぼ一致する)

このような状況を踏まえ、上述したような時系列のうち、小さいデータが何年も連続して出現することが少ないことに着目して、これらを除くことを考える。まず、 n 年間の連続した観測の時系列資料をAMS資料で整理したものを X 、そのデータを時系列順に x_i ($i=1,2,\dots,n$)とする。2年を一区切りで考える場合($m=2$),

$$Y_{21} = \{\text{Max}[x_1, x_2], \text{Max}[x_3, x_4], \dots\} \quad (6.1)$$

$$Y_{22} = \{\text{Max}[x_2, x_3], \text{Max}[x_4, x_5], \dots\} \quad (6.2)$$

という2種類の資料を抽出する。

図6.15は●が Y_{21} を示したものであり、 Y_{21} を抽出する際の区切り線もあわせて示している。○はもとのAMS資料には含まれるが、2年を一区切りとする資料には含まれないデータである。このように比較的小さい値のデータをうまく除くことができる。

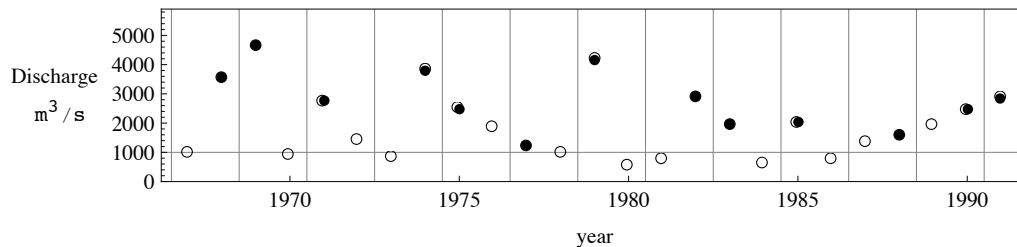


図6.15 2年で区切った洪水ピーク流量の最大値系列

このようにして求めた2つの資料 Y_{21} と Y_{22} についてAMS資料と同様に確率水文量を算出し、その平均および最大と最小の幅を求める。また、この作業を $m=3$ 以上の場合についても行う。

POT資料については、POT資料をAMS資料と比較できるようにGumbel確率紙上にプロットできるように変換したものがGumbel確率紙上で直線状を保っている下限値を閾値として設定している。(POT資料を指数確率紙にプロットした場合に直線状を保っている下限値で判断しても同じである)

6.3.2 複数年で区切って抽出した最大値系列の特性

豊川の基準点である石田地点上流の流域平均雨量104年間の年最大日雨量について複数年で区切った最大値系列の特性を調べる。

このデータをGumbel確率紙上にプロットしてみると図6.16に示すようにデータはほとんど直線状にプロットされるのでGumbel分布をあてはめるのが適当と考えられる。GEV分布をあてはめると、形状母数が0.010となり、ほぼ0であり、Gumbel分布と見なすことができる。

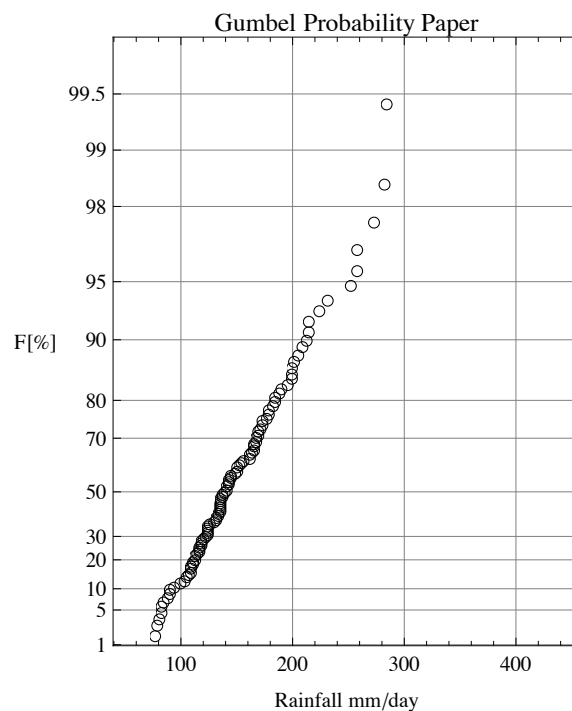


図6.16 豊川石田地点上流年最大日雨量の確率プロット

この資料を複数年で区切った最大値系列についてGEV分布をあてはめ、その形状母数を求めたものが表6.1である。区切り間隔 m が大きくなると徐々に形状母数が大きくなる傾向にある。 $m=2$ の場合はほぼGumbel分布とみなせる。なお、赤池の情報量基準AICによりGEV分布とGumbel分布を比較すると m が1から4の範囲ですべてGumbel分布の方が適合度が良い結果となった。jackknife法でリサンプリングして推定誤差を比較すると、 m が2以下でGumbel分布の方がGEV分布より推定誤差が小さくなった。 m が3と4では一部にGEVの方が推定誤差が小さいものがあるが、その他はGumbel分布の方が推定誤差が小さく、総じてGumbel分布をあてはめるべきと考えることができる。

SLSCでみると、図6.17のように m が大きくなるとSLSCが徐々に大きくなる。 m が3以上では要注意である。

表6.1 豊川石田地点上流年最大日雨量の複数年最大値資料の形状母数

m	1	2	3	4
形状母数	0.010	0.016	0.054	0.014
		-0.017	0.032	0.066
			0.036	0.107
				0.063

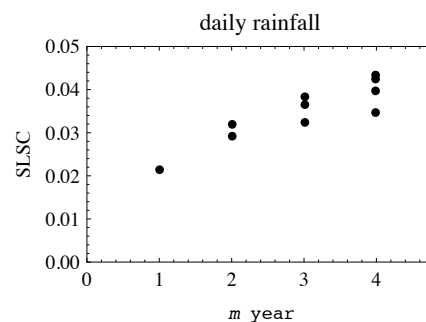


図6.17 豊川石田地点上流年最大日雨量の複数年最大値資料のSLSC

次に豊川石田地点の104年間の年最大ピーク流量についてみる。図6.18のようにこれをGumbel確率紙上にプロットすると少し上に凸の形状となる。この形状母数を求めたものを表6.2に示す。 m が1のAMS資料では裾の長い分布を示すように形状母数は負の値であるが、 m が2, 3, 4の場合はすべて正になっている。また、その値の変動も大きい。AICで見ると $m=1$ の場合、GEV分布がGumbel

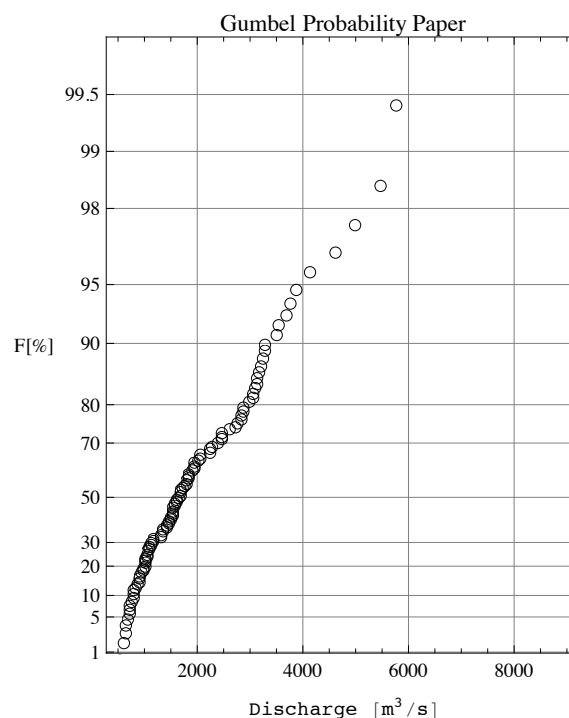


図6.18 豊川石田地点における年最大洪水ピーク流量の確率プロット

表6.2 豊川石田地点年最大洪水ピーク流量
の複数年最大値資料の形状母数

m	1	2	3	4
形状母数	-0.101	0.104	0.069	0.114
		0.021	0.027	0.054
			0.126	0.023
				0.006

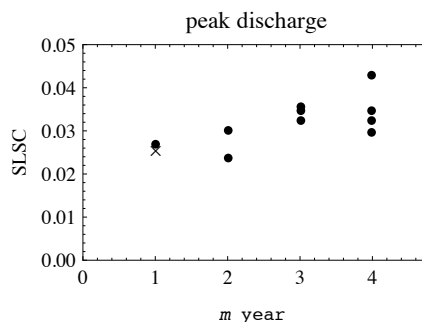


図6.19 豊川石田地点年最大洪水ピーク流量
の複数年最大値資料のSLSC

分布より適合度が高くなっているが、それ以外の場合は $m=4$ の一部を除いてGumbel分布の方が適合度が高くなっている。jickknife推定誤差は $m=1$ および2ではGumbel分布の方が小さいが、それ以上になるとGEV分布が小さい場合も混在する。Gumbel分布のSLSCは図6.19のように m が大きくなるとSLSCも徐々に大きくなる傾向にある。なお、GEV分布の $m=1$ でのSLSCを図中に×印で示してあるが、Gumbel分布のものより少し小さい状況である。

なお、ここで評価されたGEV分布は異なる複数年で区切ったそれぞれの標本に最適化された形状母数を持つものであり、形状母数が m とともに変化している。 m を変えると形状母数も変わってしまうので、複数年で区切った最大値系列に対する結果をAMS資料の結果とあわせて評価する際には注意が必要である。

6.3.3 他の河川の解析事例

まず、東海地方の雲出川の1959年から1996年までの年最大洪水ピーク流量を用いて、資料整理し、評価を行う。

図6.20にAMS資料の時系列プロットを示す。図中の水平線は $400\text{m}^3/\text{s}$ であり、POT資料の閾値である。この閾値を下回るAMS資料は5個存在しているが、 Y_{21} の最小値は $411\text{m}^3/\text{s}$ であり、これらの小さな洪水ピーク流量は含まれない。

図6.21にGumbel確率紙へのプロットを示す。Gumbel確率紙にPOT資料およびPOT資料に基づく分布モデルをプロットする際には式(4.1)を用いて変換している。この図から分かるようにAMS資料における小さい洪水データはPOT資料には含まれていない。また、最大値以外の中規模洪水は同じ超過確率に対しPOT資料がAMS資料より少し大きな流量となっている。さらに、超過確率の小さい領域ではGumbel分布とGEV分布は大きく乖離しているがExp分布とGP分布はほぼ一致している。

雲出川について $m=1$ から4について1/100確率水文量を求めた結果をPOT資料での評価結果とともに図6.22に示す。点が平均値、バーが最大値、最小値までの幅を示す。Gumbel分布は m を大きくしても確率水文量はあまり変動しない。一方、GEV分布は $m=1$ の場合と2以上の場合とで確率水文量が大き

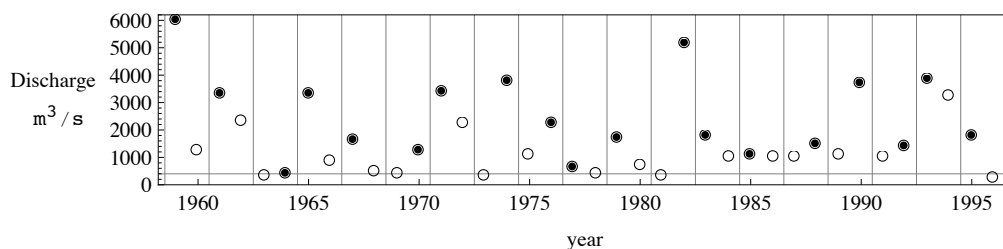


図6.20 2年で区切った雲出川の洪水ピーク流量の年最大値時系列

く異なる。 $m=2$ 以上の場合の確率水文量はGumbel分布の確率水文量に近い。 上述したように、AMS資料は規模の小さい洪水の出現状況が中規模以上のそれと異なるが、 $m=2$ とすることでこの小規模洪水は除くことができる。 この部分がなくなることにより、表6.3に示すようにGEV分布の形状母数はほぼ0または正の値を持つようになり、結果として、確率水文量が小さくなる。 逆に、Gumbel分布は小規模洪水の影響がなくなることにより、 $m=2$ で $m=1$ より少し大きめの確率水文量となっている。

POT資料の確率水文量はExp分布が約7,000m³/sでGP分布が約7,200m³/sである。 Exp分布はGumbel分布と対比できるものであるが、Exp分布は $m=1,2,3$ のGumbel分布より800ないし600m³/s大きくなっている。 確率紙に注目すると、2,000m³/s以上の洪水で年最大値に次ぐ洪水が比較的多くプロットされており、これがAMS資料に対してPOT資料の確率水文量が大きくなっている原因と考えられる。 特にPOT資料の上位2番目のデータは年最大値ではなく、これがPOT資料とAMS資料の確率紙上での乖離の大きな原因となっている。

上の事例ではGumbel分布の $m=1$ および2の確率水文量はPOT資料のExp分布とGP分布の間に挟まれていないが、東海地方では多くの場合は挟まれている。 以下に上述の事例とはパターンが異なる事例を紹介する。

まず、図6.23に示す櫛田川の確率水文量は、Gumbel分布については $m=1,2,3$ ともに同程度であり、これとGEV分布の $m=1$ がほぼ同じである。 GEV分布は m が大きくなると徐々に確率水文量が小さくなっており、 m が3以上でGP分布より小さな確率水文量を与える。 この状況は、図6.24のGumbel確率紙プロットにおいて小規模洪水の発生状況が中規模の洪水とほぼ同様の確率分布曲線上にあり、比較的大規模の洪水が上限を持つような傾向で出現していることに対応している。

次に図6.25に示す鈴鹿川の例では $m=1$ の場合のGumbel分布とGEV分布の確率水文量は大きく乖離しているが、 $m=2$ 以降はほぼ同じ値を取りながら m の増加とともに少しずつ増加している。 また m が3より大きいところではPOT資料のGP分布より大きな確率水文量となっている。 このことは、図6.26

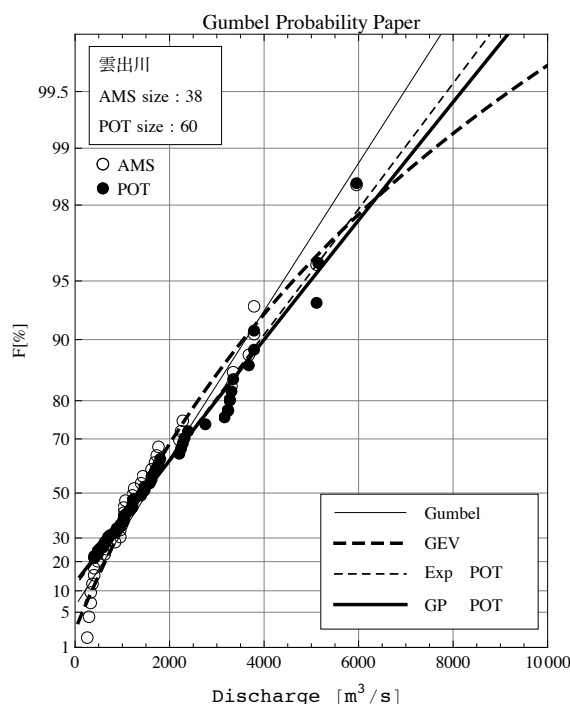


図6.21 雲出川の洪水ピーク流量のAMS資料とPOT資料の比較

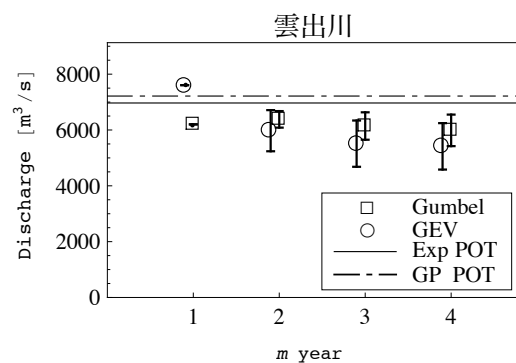


図6.22 雲出川洪水ピーク流量の複数年最大値資料による1/100確率水文量とPOT解析の比較

表6.3 雲出川洪水ピーク流量の複数年最大値資料の形状母数

m	1	2	3	4
形状母数	-0.188	-0.008	0.083	0.140
		0.250	0.321	0.629
			0.476	0.528
				0.211

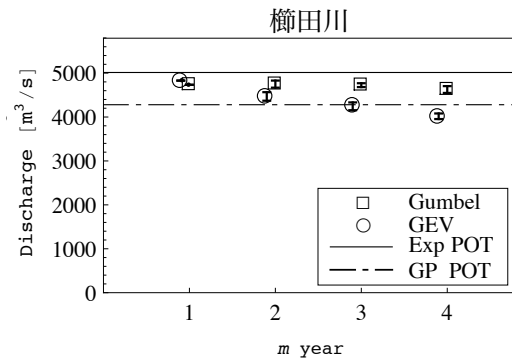


図6.23 櫛田川洪水ピーク流量の複数年最大値資料による1/100確率水文学とPOT解析の比較

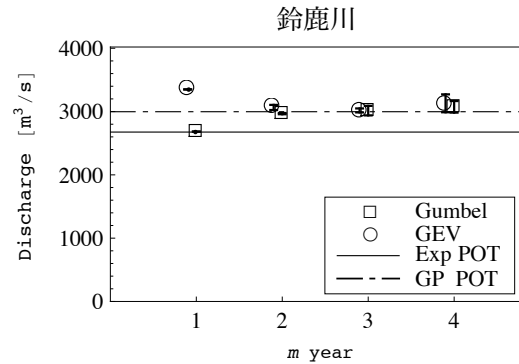


図6.25 鈴鹿川洪水ピーク流量の複数年最大値資料による1/100確率水文学とPOT解析の比較

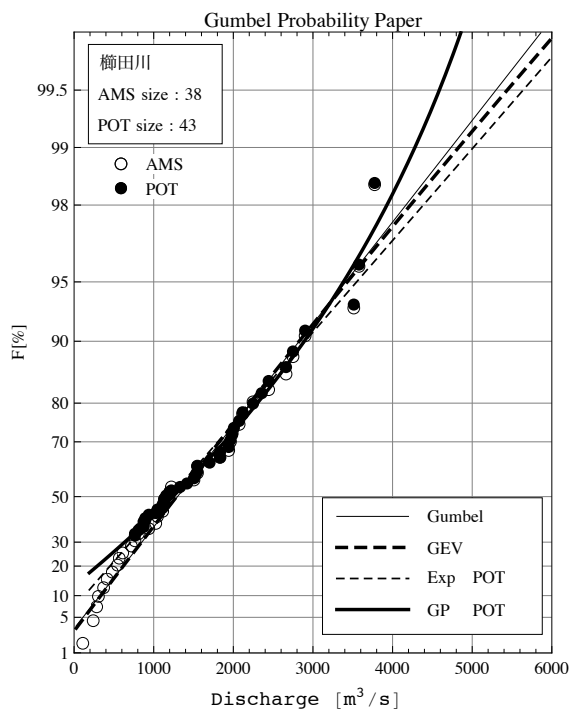


図6.24 櫛田川の洪水ピーク流量のAMSとPOTの比較

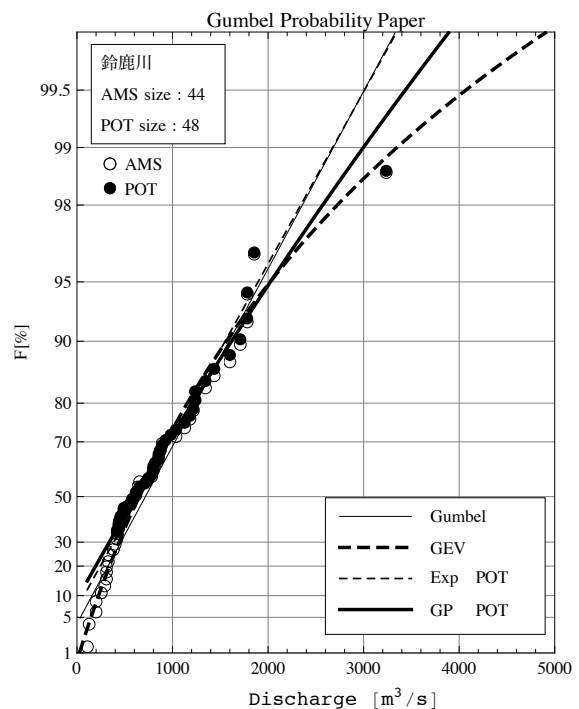


図6.26 鈴鹿川の洪水ピーク流量のAMSとPOTの比較

のGumbel確率紙へのプロットをみると500m³/s以下の小規模洪水の発生状況が500~1,500m³/sの中規模洪水とは異なる確率分布曲線上にあることと確率紙上でほぼ直線状の中規模洪水に比べ標本中の最大値である大規模洪水が分布曲線からはずれて大きいプロットとなっていることに対応している。

以上の結果を踏まえると、AMS資料を用いる場合、確率紙上で確認するとともに、確率水文学量は $m=1$ だけでなく、 $m=2$ および3などもあわせてみるべきである。雲出川の事例ではPOT資料の情報がないければ、 m を変えた場合の安定性も考慮すると確率水文学量は6,200~6,400m³/sと考えるべきであろう。この確率水文学量はPOT解析の結果である7,000m³/s~7,200m³/sとは約800m³/s異なる。図6.27は図6.21のAMS関係部分のみの確率プロットを示したものであるが、図6.21のAMS資料とPOT資料の比較プロットや複数年最大値での検討結果を踏まえてみると、 $m=1$ である通常のAMS資料では適合度や安定性のみしか判断材料はなく、たとえば、ここではGumbel分布かGEV分布かどちらかの判断を求められることになり、何が適当かの判断はなかなか難しいことがわかる。図6.28はjackknife法に

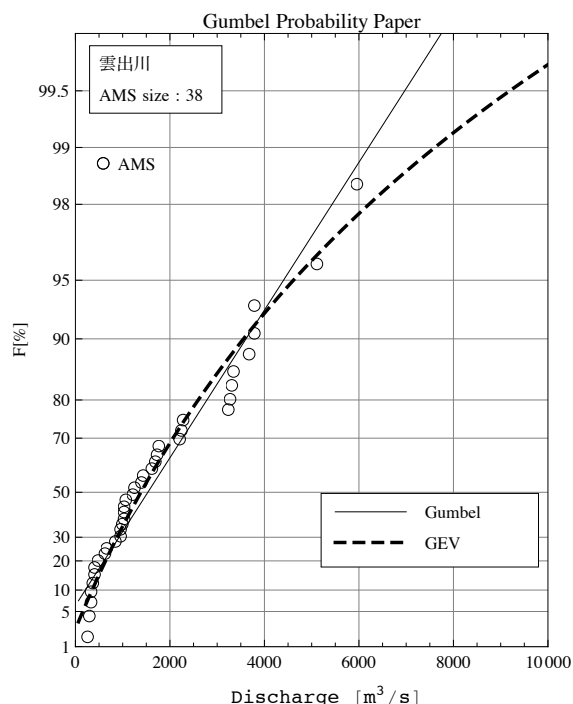


図6.27 雲出川の洪水ピーク流量のAMS資料のみでの評価

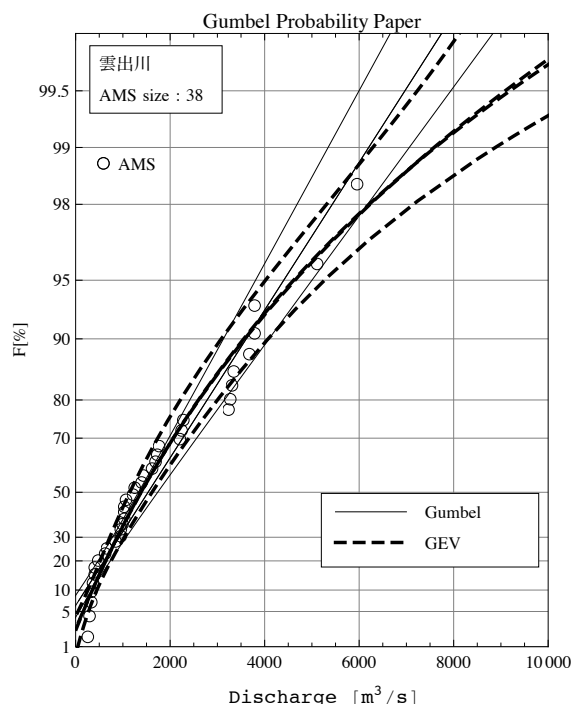


図6.28 雲出川の洪水ピーク流量のAMS資料およびjackknife法での評価

よりサンプリングの結果を示したものである。POT解析による結果である $7,000\text{m}^3/\text{s} \sim 7,200\text{m}^3/\text{s}$ はGumbel分布の推定値+推定誤差とGEV分布の推定値-推定誤差が重複している範囲の中にほぼ収まっていることが分かる。

さらに、各河川の事例をみるとPOT資料のExp分布およびGP分布で求められる確率水文量は $m=2$ のGumbel分布とかなり調和的であることがわかった。このことはAMS標本中の下位データを除く操作の中で、上位データの多くは保存されるので、POT資料と同様の確率水文量を与えるようになると考えられる。基本的に $m=2$ で下位データの影響は除かれており、 $m=3$ 以上では適合度が悪くなるなどデメリットが出てくる。しかし、雲出川のようにAMS標本中にない大きなデータがPOT資料に含まれることにより、AMS資料とPOT資料がかなり異なり、確率水文量の評価結果もかなり異なる事例もあることから、AMS資料から得られた複数年最大値だけでの判断とするのではなくPOT資料も含めて検討し、総合的な判断をするべきである。

6.3.4 まとめ

AMS資料とPOT資料をあわせて確率紙にプロットするとともに複数年を一区切りとした最大値系列を整理することにより、(1)わが国の洪水ピーク流量については非超過確率の小さい領域の事象が異なる挙動を示すことがあり、それが非超過確率の大きい確率水文量を外挿により推定する際に影響を与えることがあること、(2)その非超過確率の小さい領域の下位データが及ぼす影響を複数年を一区切りとした取り扱いをすることにより除去できる場合がかなりあり、特に、2年を一区切りとする場合が有望であること、(3)AMS資料とPOT資料の間で、第5章で示した上位一致率が低い場合、POT解析の評価をあわせて行わないと過小の確率水文量を推定する場合があること、などの知見を得た。

このような結果を踏まえ、今後は洪水ピーク資料の頻度解析においてはAMS資料だけでなくPOT資料による解析評価ができるように資料を蓄積していく必要がある。

6.4 結論

一つの時系列資料から得られたAMS資料とPOT資料を用いた解析を比較することにより、わが国の水文資料の一部にはどちらを用いてもさほど問題が起きない標本があるが、一方で、多くは3母数の分布をあてはめた時の形状母数が異なり、したがって確率水文量においても異なる結果が得られることが明らかになった。この場合、AMS解析はPOT解析に比べ大きめの確率水文量を与える。このような資料に2母数の分布をあてはめて比較すると逆にAMS解析の方がPOT解析より小さめの確率水文量を与える。このことが起こる原因はAMS資料における下位データの影響であり、これらの下位データを含まないPOT解析に優位性があることが示された。一般にAMS資料のみの解析で適合度および安定性で評価すると、Gumbel分布が選ばれる可能性が小さくないが、その場合にPOT解析より小さめの確率水文量を推定することがあるので注意する必要がある。

形状母数がAMS資料とPOT資料で異なることについて、モンテカルロシミュレーションにより検討した結果、観測期間が短いことに起因しているとははっきり確認されなかったので、データの蓄積を重ねてもこの問題は必ずしも解決されないことがわかった。

AMS解析における下位データの影響を取り除くには複数年、特に2年を一区切りとした解析が有望であることが示されたが、やはりPOT解析の結果をあわせてみる必要がある場合があることが分かった。したがって、今後はAMS解析だけでなく、POT資料による解析評価ができるように資料を蓄積していく必要がある。

参考文献

- Efron, B.: The Jackknife, the Bootstrap and Other Resampling Plans, SIAM Monograph, No.38, 92pp., 1982.
- Rasmussen, P.F., F. Ashkar, D. Rosbjerg, and B. Bobee: The POT method for flood estimation: a review, Stochastic and Statistical Methods in Hydrology and Environmental Engineering, Vol.1, pp.15-26, 1994.
- Tanaka, S. and Takara, K.: A Study on Threshold Selection in POT Analysis of Extreme Floods, The Extremes of the Extremes, IAHS Pub. no.271, pp.299-304, 2002.
- 角屋 睦：水文統計論，水工学シリーズ，土木学会水理委員会，64-02，p.59，1964.
- 宝 馨，高埴琢馬：水文頻度解析における確率分布モデルの評価規準，土木学会論文集，第393号/II-9，pp.151-160，1988.
- 田中茂信：複数年で区切った最大値資料の特性，極値理論の工学への応用，統計数理研究所共同研究リポート169，pp.1-7.，2004.
- 田中茂信，宝 馨：河川流量の頻度解析における適合度と安定性の評価，水工学論文集，第43巻，土木学会水理委員会，pp.127-132，1999.
- 田中茂信，宝 馨：洪水頻度解析におけるAMSとPDSの比較，水工学論文集，第45巻，土木学会水理委員会，pp.205-210，2001.
- 西岡昌秋，寶 馨：季節性を持つ水文時系列に基づく PDS 法と AMS 法の比較，京都大学防災研究所年報，45-B2，pp. 149-162，2002.
- 星 清：水文頻度解析，水文・水資源ハンドブック，7.3，水文・水資源学会編，朝倉書店，pp. 238-248，1997.

7. 結論

平成9年(1997)に河川法が改正され、河川計画制度が大きく変わった。それまで昭和39年(1964)に制定された新河川法に基づき河川審議会の意見を聞いて定めてられていた工事实施基本計画に代わり、河川整備基本方針および河川整備計画を定めることになった。河川整備基本方針を定める際の社会資本審議会の意見聴取、河川整備計画の案を作成しようとする場合の学識者、関係住民および関係都道府県知事又は関係市町村長の意見聴取の仕組みが取り入れられた。このように改正河川法では開かれた河川計画制度を定めているので、河川の整備計画の目標は客観性、科学的合理性をもって決められるべきである。現在、わが国の一級河川の整備目標水準は超過確率 $1/100 \sim 1/200$ であり、これまでに蓄積された時刻データに基づく水文資料の蓄積状況は平均的には50年余りであることからするとまだまだほとんどの河川は外力目標を設定する際に外挿により求めなければならない領域にある。水文統計解析の目的の一つに時系列資料から将来起こりうる確率規模相当の事象を適切に推定することがある。従来、資料整理が容易であることから年最大値資料(AMS)を用いた解析がよく用いられてきており、「水文統計解析」と言えばAMS資料を時系列資料から抽出して、これに岩井法や石原・高瀬法などの対数正規分布やGumbel分布、GEV分布などをあてはめて計画規模の確率水量を求めるというのが一般的な方法であった。多くの河川では継続的な水文観測は戦後しばらくして始まったが、一部には観測が中断した戦争前に大洪水の貴重な資料が残っている河川がある。そのデータの活用が課題となってPOT資料に関する研究をAMS資料の解析方法の研究と平行して進めることになった。このような背景から本論文ではわが国の一級水系の水文資料を対象にAMS解析とPOT解析の比較を行い、どのような水文統計解析手法を選択すべきかについて論じたものである。

各章の概要は以下のとおりである。

第1章は本論文の導入部であり、研究の背景や意義について述べた。

第2章では、わが国における河川整備と水文頻度解析の歴史について述べた。特に、既往最大洪水対応の河川整備から確率洪水規模を想定した河川整備への転換については、河川計画における大きな変化であるので、その当時の技術的背景等もあわせて整理した。また、その延長にある河川砂防技術基準の制定から改訂経緯も水文統計の記述に焦点をあてて整理した。さらに、水文極値に関する研究についてもレビューするとともに水文頻度解析に用いることができる資料集積の現状と課題を述べた。

第3章では、適合度および安定性を考慮した水文頻度解析の一般的手順および水文頻度解析によく用いられる確率分布モデルについて述べた。確率分布モデルの形状を規定する母数の推定法については、蓄積されている水文資料の観測期間が比較的短いことをふまえ、L積率法が推奨されるが、その詳細は付録に記した。従来、詳細な評価がなされてこなかった平方根指数分布および平方根指数型最大値分布(江藤ら, 1986)についてその特性を詳しく述べた。また、本研究を行うにあたって筆者がMathematicを用いて開発した水文極値の統計解析プログラムについて概要を紹介した。

第4章では、わが国の一級河川の資料を用いて年最大値系列を用いた水文頻度解析を行い、3時間、6時間、12時間、24時間の流域平均雨量資料に一般極値分布を当てはめた際の形状母数の地域変化をみるとともにL積率比(L-skewnessとL-kurtosis)の特性について調べ、形状母数が北から南へ向かうに従い負から正へ変化する傾向にあることを示した。また、洪水流量資料について適合度評価指標SLSC(Standard Least-Squares Criterion)を用いて検討し、河川流量の満足すべき適合度の基準として、 $SLSC \leq 0.04$ とするのが適当である旨提案した。また、リサンプリング手法の一つであるjackknife法を用いた安定性評価を行いその有効性を確認した。

第5章では、非毎年値(POT)資料の解析結果を毎年値(AMS)資料の結果と比較することにより、POT資料ではAMS資料に比べ形状母数が正に偏る傾向にあること、形状母数の地域内平均値は北から南へ向かうにしたがって大きくなり正に偏る傾向があること、どの地域においてもPOT資料の形状母数がAMS資料の形状母数より大きい結果となっていること、L積率比でみるとPOT資料は指数分布および一般パレート分布があてはまるものが多く、AMS資料はGumbel分布およびGEV分布があてはまるものが多いことが明らかになった。

洪水ピーク流量資料について、AMS資料と同じ標本サイズとなるように抽出したPOT資料（非毎年数最大値）は満足すべき適合度の基準として $SLSC=0.04$ をとると適合度を満足しなくなる標本が少なからずあること、適合度評価、リサンプリング手法と併せてAMS資料とPOT資料を併せて評価することにより、確率水文量の信頼性を向上させることができる等、POT資料による頻度解析が有力な手段となり得ることが明らかになった。

流域平均雨量を対象にした閾値の選定方法に関する検討では、分布モデルの母数による判断指標が良好な判断を与えること、選定された閾値を用いて抽出したPOT資料にあてはめた分布モデルは非超過確率の大きい範囲のデータに対しても良好に適合すること、AMS資料の下位データが非超過確率の大きい範囲のクォンタイルに影響を及ぼしている場合があることを明らかにした。また、わが国の水文資料を対象にする場合、POT資料による解析はAMS資料による解析に比べ優位であるので、従来のAMS資料を用いた解析に加え、POT資料による解析を行うべきである旨提案した。

第6章では、第5章で明らかにされたAMS解析とPOT解析の結果が異なることがある問題に焦点をあてた。一つの時系列資料から得られたAMS資料とPOT資料を用いた解析を比較することにより、3母数の分布をあてはめる場合、AMS解析はPOT解析に比べ大きめの確率水文量を与えるが、2母数の分布をあてはめると逆にAMS解析の方がPOT解析より小さめの確率水文量を与える。このことが起こる原因はAMS資料における下位データの影響であり、これらの下位データを含まないPOT解析に優位性があることが明らかになった。シミュレーションによりデータの蓄積を重ねても形状母数がAMS資料とPOT資料で異なる問題は残ることが分かった。この問題の解決策として、複数年、特に2年を一区切りとした解析が有望であることが分かったが、その場合でもやはりPOT解析の結果をあわせてみる必要があることが明らかになった。

わが国では、従来、資料整理の簡単なAMS解析が主に用いられてきた。角屋(1961)はAMS資料とPOT資料の理論的關係や応用例を示す中で、T年に1回起こるような事象から1年に何度も起こるような事象まで推定することが可能であり、地点雨量に基づいてAMS資料かPOT資料という問題に対してはAMS資料で実用上十分であるとしているが、本論文でみてきたように、治水計画にとって重要な流域平均雨量や洪水ピーク流量は下位データの影響によりAMS解析とPOT解析の結果が異なることが多いことが明らかになった。

近年、地球温暖化による気候変化の治水や水資源への影響が注目を集めている。POT資料はAMS資料より標本サイズを大きくとることができるので、気候変化の影響の有無を調べる際にもAMS資料より有利である。また、毎年上位数個のデータを抽出して頻度解析を行う方法も提案されている。したがって、今後は従来のAMS解析だけでなく、POT資料による解析評価や新しい解析手法が適用できるように資料を蓄積していく必要がある。

後記

1997年河川法が改正され、河川計画策定の仕組みが変わり、従来以上に計画の透明性や客観的合理性が求められるようになった。当時筆者は旧建設省から旧財団法人国土開発技術研究センターに出向しており、その仕事の中で治水計画に用いる水文統計解析についてレビューに携わることになった。最初、当センターの湧川勝己氏、岡安徹也氏とともに、山田正中央大教授、寶 馨京都大学防災研究所教授をはじめ、江藤剛治近畿大学教授、もう亡くなられた星清北海道開発局開発土木試験所所長、角屋 睦京都大学名誉教授に教を乞い、AMS資料についてのレビューを行って、商用プログラミング言語であるMathematicaによる水文統計解析プログラムを作成した。この作成過程において、当時各地方建設局の河川計画担当の技術者に無理無理使ってもらいながら、バグの発見・修正さらに使い勝手やインターフェイスなどを改良しつつなんとか使ってもらえるようなところまでこぎ着けた。なお、このソフトウェアはMathematicaのコードで記述されており、Mathematica自体がわが国の河川技術者にあまり知られていないこととそれほど廉価でないことからあまり普及することはなかった。この普及上の障害があることは重々承知していたが、Fortranでのプログラム開発の効率を考えて短期間で動くプログラムを作成するにはMathematicaが最適と考え、Mathematica上でプログラム開発することにした。なお、このプログラムはその後、財団法人国土技術研究センターのホームページ上に公開された水文統計ユーティリティーなどの基礎になっている。

このプログラム開発の過程で一部の河川では第2次世界大戦の間、観測資料がないが、戦前に大洪水を経験しており、この大洪水の資料を使わずに治水計画を検討しても沿川地域住民の理解を得ることが難しいケースがあることが分かり、これまでのAMS資料の研究とあわせてPOT解析の研究も進めることとなった。二つの解析方法を研究しているうちに、元々は一つの時系列資料から出発しているにもかかわらず、異なる確率水文量が得られることが分かり、かつ、そういうケースが少なくないことに気づいた。当時、水文統計と言えばAMS解析が主流であったことは、河川砂防技術基準（案）の記述をみればよくわかる。昭和61年版では水文統計の章の全55頁のうち、わずか2頁のみに非毎年値資料の取り扱いの解説が指数分布を中心に記述されているのみで、平成9年版では水文統計の章は16頁に縮小され、全数資料または非毎年の部分資料という表現は見られるが、非毎年値資料に関する解説部分は削除されている。わが国でPOT解析があまり用いられてこなかった背景には、理論的には同じ確率水文量を与えるから資料整理が簡単なAMS資料が用いられたと推測されるが、両者を実際のデータで比較した形跡はない。しかしながら、理論的な値や傾向と実際のデータのそれらは必ずしも一致しないことは現実には良くあることである。筆者はこの部分に注目して、POT解析の研究を進めるとともに、元々は同じ時系列資料であり同様の確率水文量を得ることができるはずと考えながらAMS解析との比較研究を行ってみた結果、AMS資料における下位データが外挿する非超過確率の大きい範囲の確率水文量に影響する問題にたどり着くことができた。POT解析における閾値選定の問題は依然解決すべき課題として残っているが、本論文では選定の目安となる方法の提案を行うとともに一つの時系列資料から抽出される極値についてAMS解析とPOT解析を併用する頻度解析手法を開発しその留意点を整理できたと考えている。

しかしながら、昨今、気候変化の影響で降雨量や洪水ピーク流量が変わると言われている。流量は人為的な活動の影響を受けるので変化することはこれまでも指摘されてきたが、IPCCの報告書に見られるように、従来、定常と考えられていた降雨量さえも変化することがかなり確実となってきた。このことをふまえ、トレンドを有するなど非定常な水文現象を対象とした頻度解析手法の研究が喫緊の課題である。

これまで、わが国の水文統計解析はAMS解析を中心に行われてきたが、解析手法としては、より多くのデータを扱うことができるPOT解析や毎年上位 r 個による解析の他、最近では分布モデルの母数

に時間変化成分を含め、その母数もあわせて同時に最尤法で解を得る方法などが開発されてきている（高橋，2008；Coles，2001）。しかしながら、水文資料の一部には時間変化が非常に大きい場合がある。このような場合に線形トレンドやベキ乗または指数型トレンドを仮定すると50年後，100年後さらに長期の極大事象は非現実的と考えられるような規模となり，これを一般に分かりやすく説明することは非常に困難である。このようなことを踏まえると，極大事象が起こりうる上限値を想定した頻度解析手法が有望になると思われる。気候変化による影響は尻上がりに増加すると考えられている。早い段階から準備すれば間に合うかもしれないが，トレンドが有意であるかないかを議論しているうちに時間は経過してしまい，トレンドがはっきり見えてきた時には既に時遅しということになる可能性がある。AMS資料のみを対象にしているとトレンドをいち早く検知することは難しい。このため，後で後悔しないよう，常に時系列の水文資料を念頭により多くのデータを用いて評価するとともにそれを極力残す努力が必要である。第2章でも触れたが，中安や米田の先進的な検討の背景には明治時代から蓄積されてきた豊富な水文資料および内外における理論的な研究があった。気候変化をふまえると非定常な現象を扱う問題意識を持ちつつデータの蓄積，解析方法の改良を続けていかなければならない。

わが国では水文統計に特化した専門書は非常に少ない。筆者の知るところでは「応用水文統計学」（岩井・石黒，1970）と「極値統計学」（Gumbel 著 河田他訳，1963）のみである。専門家に聞くと，これにはマーケットが狭いのが原因しているようである。本論文執筆に際し，際立って重要な知見は海外の文献によることが多かった。日本では角屋 睦京都大学名誉教授の一連の研究が当時世界に伍していたように思われる。というのは，”Statistics of Extremes”(Gumbel, 1958)の和訳「極値統計学」が翻訳出版されたのは米国で発刊されてから約5年後であった。この本の序にその当時のいきさつが書かれている。訳者の一人加瀬滋男が，Gumbelが近く本を出版することを聞き，出版社に翻訳を申し入れたが，個人に翻訳権は認めない中で，日本には極値理論の研究グループがあるのでそれに対してなら認めても良いということでこの本の翻訳ができたことが記されている^{注15}。この本の中で極値の第2漸近および第3漸近を扱う第7章の訳者は角屋であり，後に一般極値分布(GEV分布)に統合される対数極値A型分布および 対数極値B型分布（角屋，1956）を考案している。また，なぜかあまり知られていないが，AMS解析とPOT解析を紹介（角屋，1961）している^{注16}。このように1960年前後のわが国における水文統計分野の研究の隆盛およびその成果に比べると現在は少し寂しい感じがするのは筆者だけではないであろう。それとも水文統計は研究対象として魅力を失ってしまったのであろうか。そもそも水文統計の主な目的は将来起こるであろう未経験レベルの極端事象をこれまでの観測記録から推定することにある。定常過程でもまだ問題は山積している。そのうえ気候変化による非定常過程へと発展していかなければならない。このためには極値理論の基礎および応用に関する研究は非常に重要である。Gumbelが極値の第1漸近形式が年最大洪水量の分布関数として有用であることを指摘し，のちにこれがGumbel分布として広く実用化されてきた経緯がある（角屋，1964）。

本論文は，河川計画の水文統計に関する基準を述べたものではないので，河川計画担当者にとってどれほど役に立つかどうか分からないが，歴史的経緯を含めて極値水文統計を扱った本論文が河川管理や計画に携わる方々の参考になるとともにわが国の水文統計解析分野の活性化およびその発展に少しでも役に立つことを願いつつ結びとしたい。

注15 監訳者序によるとこの当時も採算の心配はあったようである。

注16 なお，角屋(1961)はAMS資料とPOT資料の理論的關係や応用例を示す中で，資料整理の簡単なAMS解析を行うことにより，T年に1回起こるような事象から1年に何度も起こるような事象まで推定することが可能であり，地点雨量に基づいてAMS資料かPOT資料という問題に対してはAMS資料で実用上十分であるとしている。

参考文献

- Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer, p.208, 2001.
- Gumbel, E. J., 河田龍夫, 岩井重久, 加瀬滋男監訳: グンベル 極値統計学, 生産技術センター新社, p.404, 1963. (原本"Statistics of Extremes"は1958刊行)
- 岩井重久, 石黒政儀: 応用水文統計学, 森北出版, p.370, 1970.
- 江藤剛治, 室田 明, 米谷恒春, 木下武雄: 大雨の頻度, 土木学会論文集, 第369号/II-5, pp. 165-174, 1986.
- 角屋 睦: 極値分布とその一解法, 農業土木研究, 第23巻, 第6号, pp.350-357, 1956.
- 角屋 睦: 水文統計論, 水工学シリーズ, 土木学会水理委員会, 64-02, p.59, 1964.
- 角屋 睦: 雨量分布とその年最大値の分布, 京大防災研究所年報, 第4号, pp.122-131, 1961.
- 高橋倫也: 極値統計学, 統計数理研究所, 公開講座資料, p.57, 2008.

謝辞

本研究を遂行するにあたって、京都大学防災研究所副所長の寶 馨教授には終始ご指導を頂いた。特に研究をはじめた頃、関係資料の提供や識者の紹介をはじめ統計数理研究所研究所の研究集会への参加を勧めていただいたことや国内外への論文投稿をご指導いただいたことなどあらゆる面でこの研究の推進にお世話になり心よりお礼申し上げます次第である。また、論文の完成に際し、貴重なご意見を賜るとともにご指導いただいた京都大学防災研究所岡田憲夫所長および角 哲也教授にも深甚なる謝意を表する次第である。

さらに、中央大学福岡捷二教授、山田正教授にはお会いするたびに早く論文を執筆するよう叱咤激励していただいた。現在所属している土木研究所水災害・リスクマネジメント国際センター (ICHARM) の竹内邦良センター長にはいつまでも書きそうにないのんびりした筆者に期限を設けて常にどこまでできたかと問いながら論文執筆環境を整えていただいた。ここに厚くお礼申し上げます次第である。

この研究は、全国の資料を解析して得られた知見に基づいている。膨大な資料を提供していただいた当時の建設省の河川計画担当の方々、国土交通省の方々のご協力なくしては遂行できなかったものである。ここに記して謝意を表する。

本論文の研究は私が1997年河川法が改正された年に財団法人国土開発技術研究センター（現国土技術研究センター）に出向したことが契機となっている。その当時からお世話になっている佐藤直良河川局長には研究開始直後から現在まで国土交通省における技術水準の向上・維持の重要性を常に指摘され論文執筆を叱咤激励していただいた。財団法人国土技術研究センターの湧川勝己氏、岡安徹也氏には河川計画関係のいろいろな情報を教えていただいた。また、現所属のICHARMでは、資料をまとめていただいた猪股広典氏、鈴木真理子氏、文献の収集にご尽力いただいた江橋祥子氏をはじめ皆様に大変お世話になった。これらの方々にも厚くお礼申し上げます次第である。

論文を書く心の支えを頂いたのは大学時代の恩師である野田英明先生である。また、異なる価値観の重要性を事あるごとに教えていただいた京都大学防災研究所岡田憲夫所長にも重ねてお礼を申し上げます次第である。

最後に単身赴任中、本論文を執筆する間終始支えてくれた妻尚子をはじめ家族に感謝する。

Appendix

A-1 河川工事年表

西暦	和暦	河川関連事項
323	仁徳11年	淀川末流に新川開削（現在の天満川筋）及び茨田堤修築（記録に残るわが国最初の河川工事と言われる/日本書紀）
325	仁徳13年	大和川に横野堤を築造
646	大化2年	耕地開発のため各河川に堤防築堤の詔を発す
701	大宝1年	大宝令発布（營繕令中に治水の令あり）
731	天平3年	行基による狭山池の改修
757～	天平宝字年間	天竜川に堤防が造られる
788	延暦7年	和気清麻呂による大和川付替え工事（途中で中止）
821	弘仁12年	空海による満濃池の改修
848	嘉祥1年	淀川の茨田堤を修築
1194	建久5年	荒川筋に堤防を築造（武蔵国太田庄）
1199	正治1年	揖斐川の堤防を修築（笠縫堤防）
1212	建暦2年	朝廷、幕府に命じ賀茂川堤防を修築
1247	宝治1年	利根川の堤防を修築（下河辺庄堤防）
1542	天文11年	武田信玄、釜無川に堤防を築造（のちに信玄堤と称す）
1574	天正2年	荒川筋に熊谷堤を築造
1580	天正8年	佐々成政、常願寺川に堤防を築造
1582	天正10年	豊臣秀吉、木曾川を改修
1589	天正17年	加藤清正、肥後国（熊本地方）の河川改修を始める
1592	文禄1年	豊臣秀吉、淀川を改修
1594	文禄3年	豊臣秀吉、宇治川流路を付替え堤防を築造
1596～	慶長年間	成富兵庫茂安、川上川（嘉瀬川）を改修
		細川忠興、山国川に金谷堤を築造
1597	慶長2年	安部川に堤防を築造
		足立重信、伊予川（重信川）を改修
1603	慶長8年	加藤清正、白川を坪井川と分流する
1605	慶長10年	菊地川の改修工事完成
1606	慶長11年	角倉玄之（角倉了以の子）、大堰川の航路を開削
1607	慶長12年	角倉了以、富士川・天竜川の航路を開削
1609	慶長14年	木曾川左岸堤（御囲堤）築造
		岡田将監、木曾川に水制（将監猿尾）を築造
1619	元和5年	芦田川堤防（水野土手）築造
1621	元和7年	黒田長政、栗山大膳をもちいて遠賀川に堀川開削を始める
1623	元和9年	川村孫兵衛、北上川に新川を開削し河口を石巻に遷す
1624～	寛永年間	古川善長、阿武隈川を改修し大檀（避難場）築く
		成富兵庫茂安、筑後川千栗堤防を築造
1629	寛永6年	久下村地先にて荒川の河道を入間川流路に移す
1636	寛永13年	京極忠高、斐伊川を改修
1640	寛永17年	最上川の松川筋に堤防を築造
		阿武隈川に堤防築造
1645	正保2年	岩木川及び十川の付替え
1647	正保4年	伊勢の宮川堤修築
1652～	承応年間	野中兼山、仁淀川を改修
1654	承応3年	赤堀川を開削し利根川の水を銚子に流下させる
1658	万治1年	加古川に益田堤を築造
1661～	寛文年間	雄物川の河口を新屋勝平山麓に付け替える
1669	寛文9年	旭川の放水路として百間川を設ける工事に着手
1670	寛文10年	瀬田川を浚渫
1674	延宝2年	富士川に雁堤を築造
1684	貞享1年	河村瑞賢、淀川改修（安治川）起工
1687	貞享4年	小野仙右衛門、川内川に堤防を築造
1704	宝永1年	大和川の付替え工事
1709	宝永6年	狩野川に江間堤防を築造
1726	享保11年	田中丘偶による酒匂川の改修
1730	享保15年	阿賀野川の放水路として松ヶ崎浜を開削
1754	宝暦4年	薩摩藩、木曾川改修工事に着手
1769	明和6年	常願寺川に水防林を設ける
1851	嘉永4年	田中政義、筑後川改修計画のために水理実験を行う
1872	明治5年	オランダ人技師来日
1874	明治7年	オランダ人技師らにより淀川の大阪綱島地先に粗梁水制を試設・淀川修築工事を開始
		オランダ人技師らにより木津川・宇治川の砂防工事調査始まる
1875	明治8年	オランダ人技師らにより江戸川松戸地先に粗梁水制を試設し、利根川低水工事を開始
1879	明治12年	安積疎水工事（福島県）に着手
1885	明治18年	琵琶湖疎水工事に着手
1887	明治20年	木曾川三川分流工事に着手
1896	明治29年	河川法公布
		淀川改良工事に着手（本格的な高水工事の開始）

1900	明治33年	布引ダム（五本松堰堤・神戸市）竣工（わが国最初の重力式コンクリートダム）
1904	明治37年	南郷洗堰（瀬田川、滋賀県）竣工
1909	明治42年	大河津分水工事（信濃川）に着手
1916	大正5年	芦安砂防ダム（御勅使川、山梨県）工事に着手 (本格的にコンクリートを利用した初期の砂防ダム)
1924	大正13年	荒川放水路(東京都) が完成
1932	昭和7年	太田川放水路工事（広島県）に着手
1939	昭和14年	利根川、鶴見川、那珂川、淀川の改修工事継続事業等に着手
1940	昭和15年	河水統制事業に対する補助開始 (水系で一貫した水利用を図ることが本格化)

注)

年代は、工事等の着工年・完成年などが明確に区別されていないため、他の文献と比較した場合に若干の差がある場合がある。なお、本資料の年代および項目は、江戸時代までは「明治前日本土木史（社）土木学会 昭和11年」の年表に記載されている主な項目を、明治以降は「日本土木史（社）土木学会 昭和48年」「現代日本史 高橋裕 彰国社 平成2年」「日本砂防史（社）全国治水砂防協会 昭和56年」の年表に記載されている主な項目を基本として取りまとめている。

出典) 国土交通省HP

http://www.mlit.go.jp/river/pamphlet_jirei/kasen/rekishibunka/kasengijutsu16.html

A-2 建設省河川砂防技術基準（昭和33年）の序文

序文

建設技監

工学博士 米 田 正 文

科学技術は日に日に進歩してやむところを知らない。ことに最近、工学関係の発達はめざましいものがある。

我が河川関係技術は、明治初年オランダの技術が導入されてわが国古来の技術に科学的基礎を与えられて以来今日まで、90年近い時日を経過した。

この間に、河川の治水利水技術、すなわち河川改修、砂防、ダム及び海岸に関する技術は長足の進歩をとげたといつてよいと思う。唯、この進歩がわれわれのなしうる最大のものであったかどうかには多大の疑問がある。私自身、河川関係の技術に今日までの経歴の大部分を注ぎ込んできて、反省してみると、技術進歩の方法にいろいろとくふうをこらすべきことが多かったことに痛感する。今後の技術発展のためにそれらのひとつひとつについて改善措置を講じてゆきたいというのが今日の私の念願である。

今回制定しようとする河川に関する技術基準はその現れのひとつである。

河川に関する技術は、わが国において、最も早くより発達したにもかかわらず、これに関する技術基準がほとんどなかったといつても過言ではないと思う。そのため新進の学者は先輩のたどった道を繰返すことが多かった。成功の場合でも失敗の場合でも、これは改めるべきである。新進者は少なくとも先輩の築き上げた技術の水準を基礎として、それよりの進歩に努めなければならない。かくて技術は段階的に進歩するものである。この段階を示すのがこの技術基準である。すなわち、今日の段階における建設省の河川に関する技術水準はこの基準書に示す通りであるが、来年の技術水準はこの基準書よりもさらに進歩すべきものであつて、できれば毎年この基準書は改訂を加えてゆくべき性格のものである。かようにすることによって建設省の技術基準は毎年進歩し、しかもその進歩の跡が明らかになり、新進者の研究を促進することになる。かくて河川に関する技術は秩序よく、能率的に、加速度的に前進するであろうと思う。これが私のこの技術基準書を制定しようとした動機であるが、この基準が将来の河川技術の基礎となつてけんらんたる技術の殿堂が築かれることを固く信じている。

この基準を制定するために二ヶ年余りの時日を要した。また、このために本省、土木研究所、各地方建設局のそれぞれの権威者に執筆を依頼し、それを関係者の数回の審議にかけ、練ったものである。しかしながら、約90年間に蓄積された技術を簡明に一冊の本にまとめることはなかなか至難のことであつて、この基準はじゅうぶんなる審議を尽くしたとはいえないが一応の取りまとめをし、今後さらに慎重な検討を加えて万全なものにしたいと思う。

河川に関する技術基準書の完成されるにあたり、私の所懐を述べ、本書の編集に関係された多くの方々に熱く感謝するとともに、この書の完成日を同志とともに喜びたい。

昭和33年5月20日

A-3 積率法とPWM法, L積率法

確率分布モデルを標本にあてはめる際の母数を求める方法は、PWM法およびL積率法が考案される以前は積率法が中心であった。Hosking and Wallis(1997)に積率法とPWM法, L積率法の関係が分かりやすく説明されているので紹介する。なお、わが国の水文資料の観測期間が比較的短いことをふまえるとL積率法は適している母数推定法である。

積率法においては、原点のまわりの r 次の積率 ν_r は以下のように定義される。

$$\nu_r = \int_{-\infty}^{\infty} x^r f(x) dx \quad (\text{A3.1})$$

また、平均値のまわりの r 次の積率 μ_r は次式で与えられ、

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x) dx \quad (\text{A3.2})$$

平均、分散、ひずみ係数はそれぞれ

$$\mu = \nu_1 = \int_{-\infty}^{\infty} x f(x) dx \quad (\text{A3.3})$$

$$\sigma^2 = \mu_2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \nu_2 - \mu^2 \quad (\text{A3.4})$$

$$\gamma = \mu_3 / \sigma^3 \quad (\text{A3.5})$$

であり、これらの母集団の統計量に対し、標本を用いて得られる統計量が母集団と標本でそれぞれ等しいとおいて母数を求める。なお、標本分散、標本ひずみ係数については不偏推定量を用いる必要がある。観測値を (X_1, X_2, \dots, X_n) とすると、これらは次のように表される。

$$\hat{\mu}_X = \bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (\text{A3.6})$$

$$\hat{\sigma}_X^2 = S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \quad (\text{A3.7})$$

$$\hat{\gamma}_X = G = \frac{n \sum_{i=1}^n (X_i - \bar{X})^3}{(n-1)(n-2)S^3} \quad (\text{A3.8})$$

一方、高次の積率は変数の高次のべき乗を含んでおり、これらが大きなバイアスを有することがあることがWallis et al. (1974)らにより指摘されてきた。さらに、式(A3.8)で示される $\hat{\gamma}_X$ の限界値が次式のように標本の大きさ n によって変わる。(Wilkins, 1944)

$$|\hat{\gamma}_X| \leq n^{1/2} \quad (\text{A3.9})$$

このようなことから、歪んだ分布の標本積率に基づいた推定は非常に信頼性が低いといわれている。

このような問題意識のもと、より満足のいく方法がL積率法によって示された。L積率法は、Greenwoodら(1979)により提案されたPWM法を修正したものである。

Probability Weighted Momentsは次式で定義される。

$$M_{p,r,s} = E \left[X^p \{F(X)\}^r \{1-F(X)\}^s \right] \quad (\text{A3.10})$$

ここに、 F は変量 X の累積分布関数である。この中でも特に有用なのが $p=1$ の次の場合である。

$$\alpha_r = M_{1,0,r}, \quad \beta_r = M_{1,r,0} \quad (\text{A3.11})$$

確率変数 X の関数 $g(X)$ の期待値は次のように考えることができるので、

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) dF(x) = \int_{-\infty}^{\infty} g(x) f(x) dx = \int_0^1 g(x(u)) du \quad (\text{A3.12})$$

式(A3.10)および(A3.12)より、式(A3.11)は次のように表せる。

$$\alpha_r = \int_0^1 x(u) (1-u)^r du, \quad \beta_r = \int_0^1 x(u) u^r du \quad (\text{A3.13})$$

変数の高次のベキ乗を含む通常の r 次の積率

$$E[X^r] = \int_0^1 \{x(u)\}^r du \quad (\text{A3.14})$$

と比較すると、式(A3.13)のPWMにおいては変量ではなく累積分布関数 u のベキ乗となっており、累積分布関数のベキ乗で重み付けされた期待値となっている。したがって、式(A3.10)の定義では「確率加重積率」と訳されているが、以下に示すように実際には $p=1, s=0$ の式(A3.13)を用いるので「確率の r 乗で加重された平均」と考えてもよい。

α_r あるいは β_r で表されるPWMはLandwehr, Matalas and Wallis (1979a,b), Greis and Wood (1981), Wallis (1981, 1982), Hosking, Wallis and Wood (1985a)およびHosking and Wallis (1987)によって確率分布の母数を推定する際に用いられてきた。しかし、これらは確率分布のscale (尺度) やshape (形状) をみるものさしとしては直観的に理解するのは困難である。たとえば、尺度母数の推定量は $\alpha_0 - 2\alpha_1$ あるいは $2\beta_1 - \beta_0$ であり(Landwehr, Matalas and Wallis 1979a; Hosking, Wallis and Wood 1985b), 歪み係数については $6\beta_2 - 6\beta_1 + \beta_0$ (Stedinger, 1983) というようになっている。このように尺度母数および形状母数をPWMの線形結合で表すには $(1-u)^r$ や u^r の多項式で加重するより以下に示すような直交多項式を用いて加重して積分するほうが自然である。L積率は次式で表されるShifted Legendre Polynomials

$$P_r^*(u) = \sum_{k=0}^r P_{r,k}^* u^k \quad (\text{A3.15})$$

を用いて次式のようにすっきりした形に定義される.

$$\lambda_r = \int_0^1 x(u) P_{r-1}^*(u) du \quad (\text{A3.16})$$

ここに,

$$P_{r,k}^* = (-1)^{r-k} \binom{r}{k} \binom{k+r}{k} = \frac{(-1)^{r-k} (k+r)!}{(k!)^2 (r-k)!} \quad (\text{A3.17})$$

である. 4次までのL積率とPWMの関係は以下ようになる.

$$\lambda_1 = \alpha_0 = \beta_0 \quad (\text{A3.18})$$

$$\lambda_2 = \alpha_0 - 2\alpha_1 = 2\beta_1 - \beta_0 \quad (\text{A3.19})$$

$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 = 6\beta_2 - 6\beta_1 + \beta_0 \quad (\text{A3.20})$$

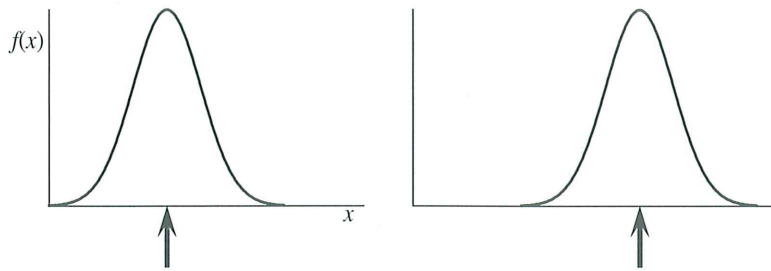
$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 = 20\beta_3 - 30\beta_2 + 12\beta_1 + \beta_0 \quad (\text{A3.21})$$

また, L積率比を次式で定義する.

$$\tau = \lambda_2 / \lambda_1 \quad (\text{L-CV}) \quad (\text{A3.22})$$

$$\tau_r = \lambda_r / \lambda_2, \quad r = 3, 4, \dots \quad (\text{A3.23})$$

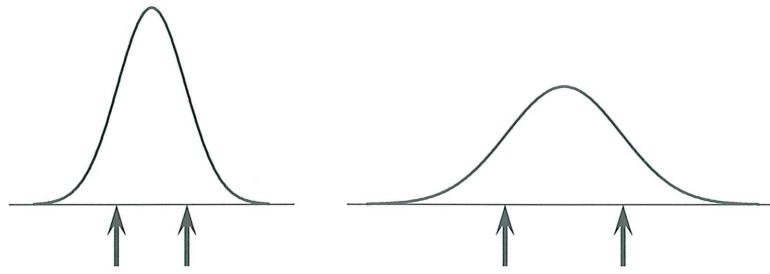
L積率の具体的な意味を順序統計量との関係でみる. まず, 小さい方から順に並べた順序統計量を $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ とし, この標本から抽出された小標本の形状を考える. 標本サイズ1の場合, $X_{1:1}$ の値しかない. これは分布の位置情報, 即ち変量の大きさ (平均値に相当する) を示している. (図A3.1参照)



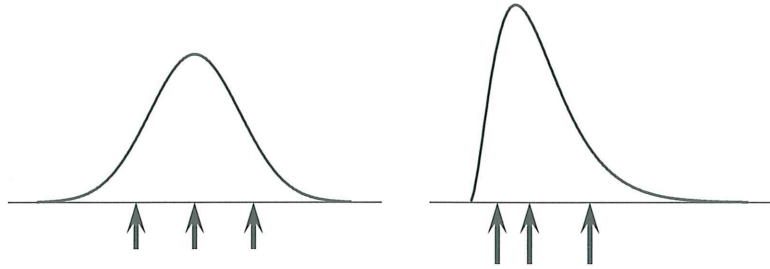
図A3.1 1次のL積率定義の概念

次に, 標本サイズ2の場合, 観測値は2つあり, 図A3.2に示すように, これらの値は分布の広がり (分散) を示していると考えられ, これらの差 $X_{2:2} - X_{1:2}$ が尺度の物差となる.

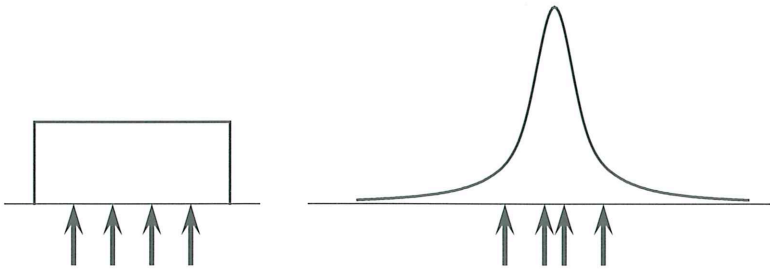
標本サイズ3の場合, 3つの値は分布の歪みを表す. 図A3.3に示すように, 分布が中央の値に対して対称ならば, 右側の2つの値の差 $X_{3:3} - X_{2:3}$ と左側の差 $X_{2:3} - X_{1:3}$ は等しくなる. したがって, これらの差の差 $X_{3:3} - 2X_{2:3} + X_{1:3}$ が0になる. もし, 右に尾をひく分布であれば, 図の右のように $X_{3:3} - X_{2:3}$ の方



図A3.2 2次のL積率定義の概念図



図A3.3 3次のL積率定義の概念図



図A3.4 4次のL積率定義の概念図

が $X_{2:3} - X_{1:3}$ より大きくなり、 $X_{3:3} - 2X_{2:3} + X_{1:3}$ は正の値となる。逆に負であれば左に尾をひく分布となる。このように $X_{3:3} - 2X_{2:3} + X_{1:3}$ は分布の歪み度の物差となる。

標本サイズ 4 の場合も同様に隣あう順序統計量の差を考え、同様な操作を行うと $X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}$ が分布の尖り度を表す物差となることが分かる。 $X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}$ の値は、図 A3.4 に示すように、一樣な分布では 0 で、尖っている分布では大きな値となる。

このように順序統計量の要素の上述の線形組み合わせは、元の標本の位置、尺度、形状に関する情報を有している。L 積率は、これらの線形組み合わせの期待値として定義される。L 積率の”L”は L 積率が順序統計量の線形組み合わせにより構成されていることを示している。L 積率は次のように定義されており、一般形は式(A3.28)で示される。

$$\lambda_1 = E(X_{1:1}) \quad (\text{A3.24})$$

$$\lambda_2 = \frac{1}{2} E(X_{2:2} - X_{1:2}) \quad (\text{A3.25})$$

$$\lambda_3 = \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3}) \quad (\text{A3.26})$$

$$\lambda_4 = \frac{1}{4} E(X_{4:4} - 3 X_{3:4} + 3 X_{2:4} - X_{1:4}) \quad (\text{A3.27})$$

$$\lambda_r = \frac{1}{r} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} E(X_{r-j:r}) \quad (\text{A3.28})$$

L積率は以下のような特性を有している。まず、分布の平均が存在すれば、すべてのL積率が存在し、L積率は1つの分布が決まり、同じL積率を有する分布は存在しない。各次数のL積率およびL積率比は次のように呼ばれる。 λ_1 : L-location (平均), λ_2 : L-scale, τ : L-CV, τ_3 : L-skewness, τ_4 : L-kurtosisであり、それぞれの値のとりうる範囲は、

λ_1 : どんな値も有りうる。

λ_2 : 0 以上,

τ : 変量Xが正の値の分布に対し, $0 \leq \tau < 1$,

3次以上のL積率比 に対して, $|\tau_r| < 1$ であり,

τ_3 と τ_4 の間には次の関係がある。

$$\frac{1}{4} (5 \tau_3^2 - 1) \leq \tau_4 < 1 \quad (\text{A3.29})$$

また、正の値の分布では τ と τ_3 は $2\tau - 1 \leq \tau_3 < 1$ の関係がある。

さらに、変量XおよびYのL積率がそれぞれ λ_r および λ_r^* であるとき、 $Y = aX + b$ であれば、次のように線形性が保たれている。

$$\lambda_1^* = a \lambda_1 + b \quad (\text{A3.30})$$

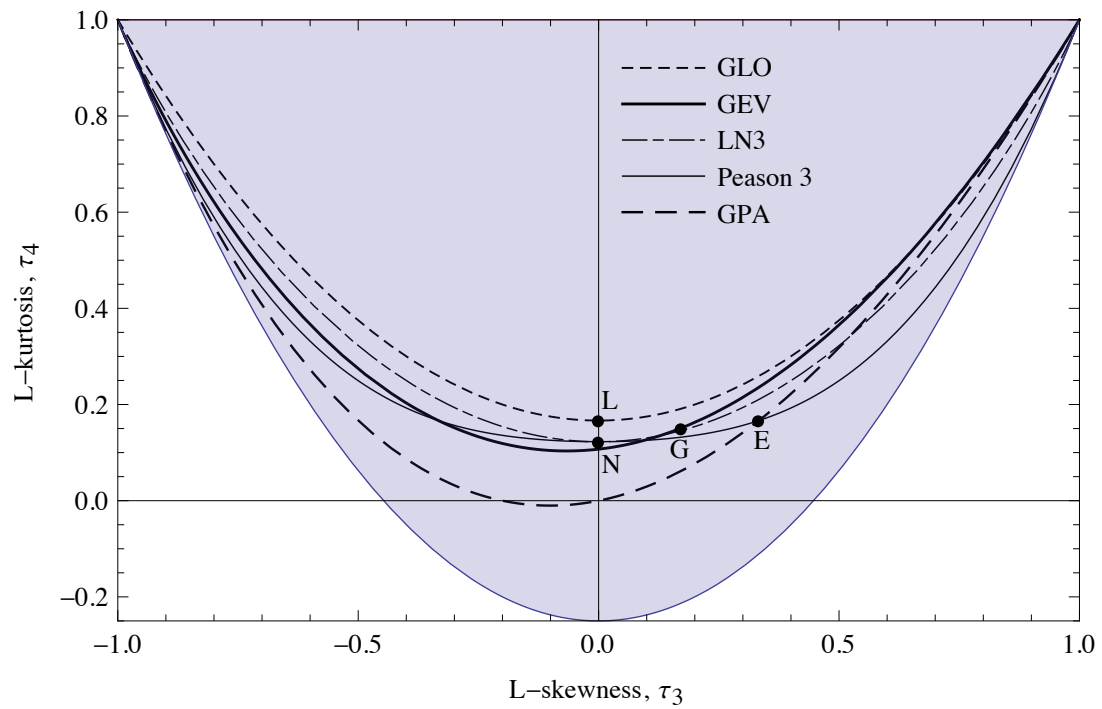
$$\lambda_2^* = |a| \lambda_2 \quad (\text{A3.31})$$

$$\tau_r^* = (\text{sign } a)^r \tau_r, \quad r \geq 3 \quad (\text{A3.32})$$

もし、Xが平均値のまわりに対象な分布であれば、3次以上の奇数次のL積率比 τ_r は0となる。L積率 λ_1 , λ_2 , τ , L積率比 τ_3 , τ_4 は確率分布を把握するのに有用である。

図A3.5は種々の分布のL-skewness τ_3 とL-kurtosis τ_4 を比較したものである。2母数の分布 (E : 指数分布, G : Gumbel分布, L : Logistic分布, N : 正規分布) は母数が変化してもL-locationとL-scaleが変わるのみで3次以上のL積率は一定値であるので、図中では点で示されている。また、対数正規分布(LN3)とPearson type III分布(Pearson 3)は $\tau_3=0$ で正規分布(N)と重なる。

従来、わが国ではL積率法 (PWM法) の積率法に対するバイアスや誤差分散が小さいことが評価されてきているが、さらなる利点として、積率法では式(A3.9)の標本歪み係数のとりうる限界値が標本サイズによって変化するので標本サイズに関係せずに種々の分布モデルを比較することができないのに対し、L積率法では図A3.5のように種々の分布モデルのL-skewnessとL-kurtosisが一枚の図に表されているので、標本のL-skewness τ_3 とL-kurtosis τ_4 を知れば、概ねどの分布に近いかが判断できる利点がある。



図A3.5 各種分布のL-skewnessとL-kurtosisの比較

2母数の分布は点で、3母数の分布は線で示されている。着色部分は式(3.29)で表される3次および4次のL積率の存在する範囲である。各記号は、それぞれ、E：指数分布、G：Gumbel分布、L：Logistic分布、N：正規分布、GLO：generalized logistic分布、GEV：一般極値分布、GPA：一般パレート分布、LN3：3母数対数正規分布、Pearson 3：Pearson type III 分布を示す。

A-4 よく用いられる確率分布モデルのL積率解

Hosking and Wallis (1997)が著した”Regional Frequency Analysis”から，水文頻度解析でよく用いられる確率分布モデルの母数，変数 x の分布範囲，確率密度関数，累積分布関数，L積率およびL積率比を示すとともにその母数推定法について述べる．

1) 指数分布

2母数：分布の下限 ξ ，尺度 α

x の分布範囲： $\xi \leq x < \infty$

$$\text{確率密度関数} : f(x) = \frac{\exp\left\{-\frac{x-\xi}{\alpha}\right\}}{\alpha} \quad (\text{A4.1})$$

$$\text{累積分布関数} : F(x) = 1 - \exp\left\{-\frac{x-\xi}{\alpha}\right\} \quad (\text{A4.2})$$

$$\text{クオンタイル} : x(F) = \xi - \alpha \log(1 - F) \quad (\text{A4.3})$$

$$\text{L積率} : \quad \lambda_1 = \xi + \alpha, \quad \lambda_2 = \frac{1}{2} \alpha, \quad \tau_3 = \frac{1}{3}, \quad \tau_4 = \frac{1}{6} \quad (\text{A4.4})$$

$$\text{母数推定法} : \alpha = 2 \lambda_2, \quad \xi = \lambda_1 - \alpha \quad (\text{A4.5})$$

なお，分布の下限 ξ が既知の場合， $\alpha = \lambda_1 - \xi$ で求められ，L積率法，積率法，最尤法は同一となる．

2) Gumbel分布

2母数：位置 ξ ，尺度 α

x の分布範囲： $-\infty \leq x < \infty$

$$\text{確率密度関数} : f(x) = \frac{\exp\left\{-\frac{x-\xi}{\alpha}\right\}}{\alpha} \exp\left[-\exp\left\{-\frac{x-\xi}{\alpha}\right\}\right] \quad (\text{A4.6})$$

$$\text{累積分布関数} : F(x) = \exp\left[-\exp\left\{-\frac{x-\xi}{\alpha}\right\}\right] \quad (\text{A4.7})$$

$$\text{クオンタイル} : x(F) = \xi - \alpha \log(-\log F) \quad (\text{A4.8})$$

$$\begin{aligned} \text{L積率} : \quad \lambda_1 &= \xi + \alpha\gamma, \quad \gamma = 0.5772 \dots \\ \lambda_2 &= \alpha \log 2, \end{aligned}$$

$$\tau_3 = 0.1699 = \log\left(\frac{9}{8}\right), \quad \tau_4 = 0.1504 = \frac{16 \log 2 - 10 \log 3}{\log 2} \quad (\text{A4.9})$$

$$\text{母数推定法: } \alpha = \lambda_2 / \log 2, \quad \xi = \lambda_1 - \alpha \gamma \quad (\text{A4.10})$$

3) 正規分布

正規分布は水文極値資料の頻度解析において必ずしもよく使われるものではないが、他の分布の形状母数が特殊な場合に正規分布になることからここで紹介する。

2母数：位置 μ , 尺度 σ

x の分布範囲： $-\infty \leq x \leq \infty$

$$\text{確率密度関数: } f(x) = \frac{1}{\sqrt{2\pi} \sigma} \exp\left\{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2\right\} = \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \quad (\text{A4.11})$$

$$\text{累積分布関数: } F(x) = \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^x \exp\left\{-\frac{1}{2} \left(\frac{t-\mu}{\sigma}\right)^2\right\} dt = \Phi\left(\frac{x-\mu}{\sigma}\right) \quad (\text{A4.12})$$

$$\text{ここに, } \phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2\right), \quad \Phi(x) = \int_{-\infty}^x \phi(t) dt \quad (\text{A4.13})$$

クオンタイル：陽的な解はない

$$\text{L積率: } \lambda_1 = \mu, \quad \lambda_2 = 0.5642 \sigma = \sigma / \sqrt{\pi}$$

$$\tau_3 = 0, \quad \tau_4 = 0.1226 = \frac{30}{\pi} \arctan \sqrt{2} - 9 \quad (\text{A4.14})$$

$$\text{母数推定法: } \mu = \lambda_1, \quad \sigma = \sqrt{\pi} \lambda_2 \quad (\text{A4.15})$$

4) 一般パレート分布

3母数：位置 ξ , 尺度 α , 形状 k

x の分布範囲： $\xi \leq x \leq \xi + \alpha/k$ ($k > 0$), $\xi \leq x < \infty$ ($k \leq 0$)

$$\text{確率密度関数: } f(x) = \frac{1}{\alpha} \left\{1 - k \frac{x-\xi}{\alpha}\right\}^{1/k-1} = \frac{1}{\alpha} \exp\{-(1-k)y\},$$

$$y = \begin{cases} -k^{-1} \log \{1 - k(x-\xi)/\alpha\} & k \neq 0 \\ (x-\xi)/\alpha & k = 0 \end{cases} \quad (\text{A4.16})$$

$$\text{累積分布関数: } F(x) = 1 - \left(1 - k \frac{x-\xi}{\alpha}\right)^{1/k} = 1 - \exp(-y) \quad (\text{A4.17})$$

$$\text{クオンタイル} : x(F) = \begin{cases} \xi + \alpha \left\{ 1 - (1-F)^k \right\} / k & k \neq 0 \\ \xi - \alpha \log(1-F) & k = 0 \end{cases} \quad (\text{A4.18})$$

$$\text{L積率} : \quad \lambda_1 = \xi + \alpha / (1+k), \quad \lambda_2 = \frac{\alpha}{(1+k)(2+k)}$$

$$\tau_3 = \frac{1-k}{3+k}, \quad \tau_4 = \frac{(1-k)(2-k)}{(3+k)(4+k)} \quad (\text{A4.19})$$

$$\tau_4 = \frac{\tau_3(1+5\tau_3)}{5+\tau_3} \quad (\text{A4.20})$$

母数推定法：下限値 ξ が既知の場合

$$k = \frac{(\lambda_1 - \xi)}{\lambda_2} - 2, \quad \alpha = (1+k)(\lambda_1 - \xi) \quad (\text{A4.21})$$

下限値 ξ が未知の場合

$$k = \frac{1-3\tau_3}{1+\tau_3}, \quad \alpha = (1+k)(2+k)\lambda_2, \quad \xi = \lambda_1 - (2+k)\lambda_2 \quad (\text{A4.22})$$

5) 一般極値分布(GEV)

3 母数：位置 ξ , 尺度 α , 形状 k

$$x \text{ の分布範囲} : \begin{cases} -\infty < x \leq \xi + \alpha/k & (k > 0) \\ -\infty < x < \infty & (k = 0) \\ \xi + \alpha/k \leq x < \infty & (k < 0) \end{cases}$$

$$\text{確率密度関数} : f(x) = \frac{1}{\alpha} \left(1 - \frac{k(x-\xi)}{\alpha} \right)^{\frac{1}{k}-1} \exp \left\{ - \left(1 - \frac{k(x-\xi)}{\alpha} \right)^{1/k} \right\}$$

$$f(x) = \frac{1}{\alpha} \exp \{ -(1-k)y - \exp(-y) \},$$

$$y = \begin{cases} -k^{-1} \log \{ 1 - k(x-\xi)/\alpha \} & k \neq 0 \\ (x-\xi)/\alpha & k = 0 \end{cases} \quad (\text{A4.23})$$

$$\text{累積分布関数} : F(x) = \exp \left[- \left(1 - k \frac{x-\xi}{\alpha} \right)^{1/k} \right] = \exp \{ -\exp(-y) \} \quad (\text{A4.24})$$

$$\text{クオンタイル} : x(F) = \begin{cases} \xi + \alpha \{1 - (-\log F)^k\} / k & (k \neq 0) \\ \xi - \alpha \log(-\log F) & (k = 0) \end{cases} \quad (\text{A4.25})$$

L積率： $k > -1$ の場合のみ定義される。

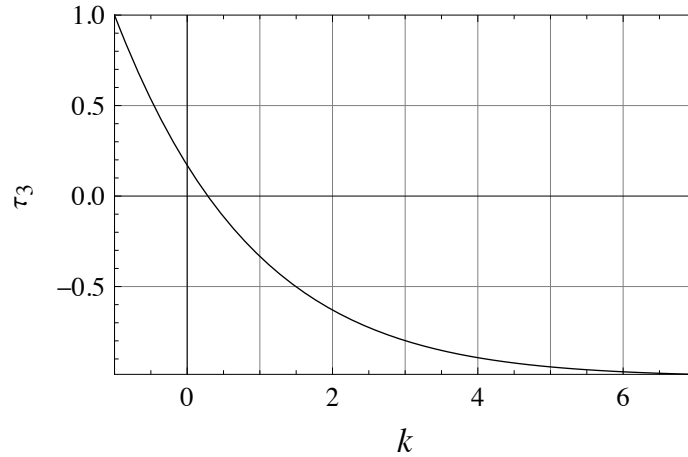
$$\lambda_1 = \xi + \alpha \{1 - \Gamma(1+k)\} / k \quad (\text{A4.26})$$

$$\lambda_2 = \alpha (1 - 2^{-k}) \Gamma(1+k) / k \quad (\text{A4.27})$$

$$\tau_3 = 2 (1 - 3^{-k}) / (1 - 2^{-k}) - 3 \quad (\text{A4.28})$$

$$\tau_4 = \{5 (1 - 4^{-k}) - 10 (1 - 3^{-k}) + 6 (1 - 2^{-k})\} / (1 - 2^{-k}) \quad (\text{A4.29})$$

ここに、 $\Gamma(\cdot)$ はガンマ関数である。式(A4.28)の関係を図A4.1に示す。



図A4.1 GEV分布の形状母数 k と τ_3 の関係（式A4.28）

母数推定法：式(A4.28)に対して、陽的な解がないので次の近似式を使う。

$$k \approx 7.8590 c + 2.9554 c^2, \quad c = \frac{2}{3 + \tau_3} - \frac{\log 2}{\log 3} \quad (\text{A4.30})$$

$$\alpha = \frac{\lambda_2 k}{(1 - 2^{-k}) \Gamma(1+k)}, \quad \xi = \lambda_1 - \alpha \{1 - \Gamma(1+k)\} / k \quad \text{注}^{17} \quad (\text{A4.31})$$

6) generalized logistic分布(GLO)

3母数：位置 ξ ，尺度 α ，形状 k ，（ $k=0$ の場合はlogistic分布）

^{注17} Handbook of Hydrologyのp.18.18, 式(18.2.22c)は間違っている。Table18.1.2および本文の式(A4.31)が正しい。

$$x \text{ の分布範囲 : } \begin{cases} -\infty < x \leq \xi + \alpha/k & (k > 0) \\ -\infty < x < \infty & (k = 0) \\ \xi + \alpha/k \leq x < \infty & (k < 0) \end{cases}$$

$$\text{確率密度関数 : } f(x) = \frac{\alpha^{-1} \exp\{-(1-k)y\}}{(1 + \exp(-y))^2},$$

$$y = \begin{cases} -k^{-1} \log\{1 - k(x - \xi)/\alpha\} & k \neq 0 \\ (x - \xi)/\alpha & k = 0 \end{cases} \quad (\text{A4.32})$$

$$\text{累積分布関数 : } F(x) = 1 / \{1 + \exp(-y)\} \quad (\text{A4.33})$$

$$\text{クオンタイル : } x(F) = \begin{cases} \xi + \alpha \{1 - \{(1 - F)/F\}^k\} / k & k \neq 0 \\ \xi - \alpha \log\{(1 - F)/F\} & k = 0 \end{cases} \quad (\text{A4.34})$$

L積率： $-1 < k < 1$ の場合のみ定義される。

$$\lambda_1 = \xi + \alpha(1/k - \pi/\sin k\pi) \quad (\text{A4.35})$$

$$\lambda_2 = \alpha k \pi / \sin k\pi \quad (\text{A4.36})$$

$$\tau_3 = -k \quad (\text{A4.37})$$

$$\tau_4 = (1 + 5k^2) / 6 \quad (\text{A4.38})$$

$$\text{母数推定法 : } k = -\tau_3, \quad \alpha = \frac{\lambda_2 \sin k\pi}{k\pi}, \quad \xi = \lambda_1 - \alpha \left(\frac{1}{k} - \frac{\pi}{\sin k\pi} \right) \quad (\text{A4.39})$$

7) 対数正規分布

3 母数： 位置 ξ , 尺度 α , 形状 k , ($k=0$ の場合は正規分布)

$$x \text{ の分布範囲 : } \begin{cases} -\infty < x \leq \xi + \alpha/k & (k > 0) \\ -\infty < x < \infty & (k = 0) \\ \xi + \alpha/k \leq x < \infty & (k < 0) \end{cases}$$

$$\text{確率密度関数 : } f(x) = \frac{\exp\{k y - y^2/2\}}{\alpha \sqrt{2\pi}},$$

$$y = \begin{cases} -k^{-1} \log\{1 - k(x - \xi)/\alpha\} & k \neq 0 \\ (x - \xi)/\alpha & k = 0 \end{cases} \quad (\text{A4.40})$$

$$\text{累積分布関数 : } F(x) = \Phi(y) \quad (\text{A4.41})$$

ここに、 Φ は標準正規分布の分布関数であり、式(A4.13)で表される。

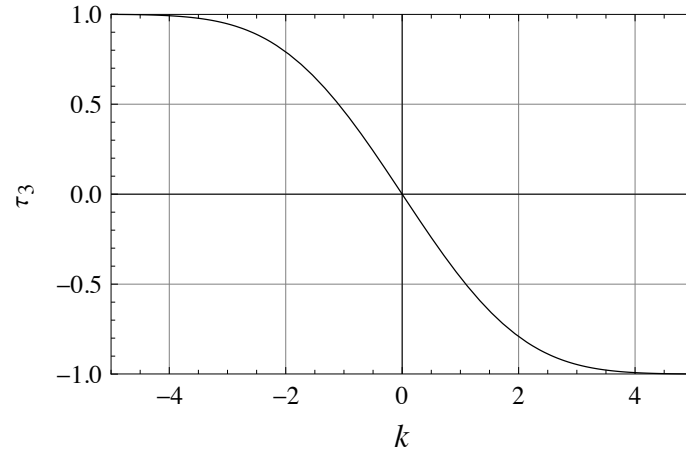
クオンタイル：陽的な解はない。

$$\text{L積率：} \quad \lambda_1 = \xi + \alpha \left\{ 1 - \exp\left(k^2 / 2\right) \right\} / k \quad (\text{A4.42})$$

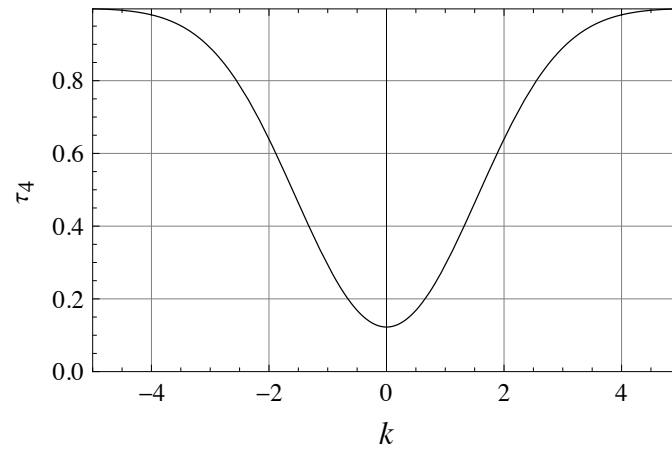
$$\lambda_2 = \frac{\alpha}{k} \exp\left(k^2 / 2\right) \left\{ 1 - 2 \Phi(-k / \sqrt{2}) \right\} \quad (\text{A4.43})$$

$$\tau_3 \approx -k \frac{A_0 + A_1 k^2 + A_2 k^4 + A_3 k^6}{1 + B_1 k^2 + B_2 k^4 + B_3 k^6} \quad (\text{A4.44})$$

$$\tau_4 \approx \tau_4^0 + k^2 \frac{C_0 + C_1 k^2 + C_2 k^4 + C_3 k^6}{1 + D_1 k^2 + D_2 k^4 + D_3 k^6} \quad (\text{A4.45})$$



図A4.2 対数正規分布の形状母数 k と τ_3 の関係（式A4.44）



図A4.3 対数正規分布の形状母数 k と τ_4 の関係（式A4.45）

母数推定法： τ_3 を用いて次式で形状母数 k を求める。

$$k \approx -\tau_3 \frac{E_0 + E_1 \tau_3^2 + E_2 \tau_3^4 + E_3 \tau_3^6}{1 + F_1 \tau_3^2 + F_2 \tau_3^4 + F_3 \tau_3^6}, \text{ for } |\tau_3| \leq 0.94, |k| \leq 3 \quad (\text{A4.46})$$

$$\alpha = \frac{\lambda_2 k \exp(-k^2/2)}{1 - 2 \Phi(-k/\sqrt{2})}, \quad \xi = \lambda_1 - \frac{\alpha}{k} \left\{ 1 - \exp(k^2/2) \right\} \quad (\text{A4.47})$$

表A4.1 式(A4.44~A4.46)の係数

$\tau_4^0 = 1.2260172 \times 10^{-1}$		
$A_0 = 4.8860251 \times 10^{-1}$	$C_0 = 1.8756590 \times 10^{-1}$	$E_0 = 2.0466534$
$A_1 = 4.4493076 \times 10^{-3}$	$C_1 = -2.5352147 \times 10^{-3}$	$E_1 = -3.6544371$
$A_2 = 8.8027039 \times 10^{-4}$	$C_2 = 2.6995102 \times 10^{-4}$	$E_2 = 1.8396733$
$A_3 = 1.1507084 \times 10^{-6}$	$C_3 = -1.8446680 \times 10^{-6}$	$E_3 = -0.20360244$
$B_1 = 6.4662924 \times 10^{-2}$	$D_1 = 8.2325617 \times 10^{-2}$	$F_1 = -2.0182173$
$B_2 = 3.3090406 \times 10^{-3}$	$D_2 = 4.2681448 \times 10^{-3}$	$F_2 = 1.2420401$
$B_3 = 7.4290680 \times 10^{-5}$	$D_3 = 1.1653690 \times 10^{-4}$	$F_3 = -0.21741801$

8) Pearson type III 分布

3 母数：位置 μ , 尺度 σ , 形状 γ

形状母数 γ が 0 でなければ次の変数変換をおこなうと, γ の正負毎に x の分布範囲, 確率密度関数, 累積分布関数が以下のように表される.

$$\alpha = \frac{4}{\gamma^2}, \quad \beta = \frac{1}{2} \sigma |\gamma|, \quad \xi = \mu - 2 \sigma / \gamma, \quad \gamma \neq 0 \quad (\text{A4.48})$$

$\gamma > 0$ の場合,

x の分布範囲: $\xi \leq x < \infty$

$$\text{確率密度関数: } f(x) = \frac{(x - \xi)^{\alpha-1} \exp\{-(x - \xi)/\beta\}}{\beta^\alpha \Gamma(\alpha)} \quad (\text{A4.49})$$

$$\text{累積分布関数: } F(x) = G\left(\alpha, \frac{x - \xi}{\beta}\right) / \Gamma(\alpha) \quad (\text{A4.50})$$

$\gamma < 0$ の場合,

x の分布範囲: $-\infty < x < \xi$

$$\text{確率密度関数: } f(x) = \frac{(\xi - x)^{\alpha-1} \exp\{-(\xi - x)/\beta\}}{\beta^\alpha \Gamma(\alpha)} \quad (\text{A4.51})$$

$$\text{累積分布関数} : F(x) = 1 - G\left(\alpha, \frac{\xi - x}{\beta}\right) / \Gamma(\alpha) \quad (\text{A4.52})$$

ここに, $G(\cdot)$ は不完全ガンマ関数

$$G(\alpha, x) = \int_0^x t^{\alpha-1} \exp(-t) dt \quad (\text{A4.53})$$

である.

$\gamma = 0$ の場合, 分布は正規分布となり, x の分布範囲, 確率密度関数, 累積分布関数が以下のように表される.

x の分布範囲: $-\infty < x < \infty$

$$\text{確率密度関数} : f(x) = \phi\left(\frac{x - \mu}{\sigma}\right) \quad (\text{A4.54})$$

$$\text{累積分布関数} : F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right) \quad (\text{A4.55})$$

Pearson type III 分布は形状母数 $\gamma=2$ の場合, 指数分布, $\gamma=-2$ の場合, 逆指数分布となる.

クオンタイル: 解析的な解はない.

$$\text{L積率:} \quad \lambda_1 = \xi + \alpha\beta, \quad \lambda_2 = \frac{\beta \Gamma\left(\alpha + \frac{1}{2}\right)}{\sqrt{\pi} \Gamma(\alpha)}, \quad (\text{A4.56})$$

$$\tau_3 = 6 I_{1/3}(\alpha, 2\alpha) - 3 \quad (\text{A4.57})$$

ここに,

$$I_x(p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_0^x t^{p-1} (1-t)^{q-1} dt \quad (\text{A4.58})$$

は正則不完全ベータ関数である. τ_4 については簡単には求められない. α と τ_3 の関係および α と τ_4 の関係の近似式が式(A4.59)～式(A4.62)のように求められており, 精度は 10^{-6} である.

$\alpha \geq 1$ の場合,

$$\tau_3 \approx \alpha^{-1/2} \frac{A_0 + A_1 \alpha^{-1} + A_2 \alpha^{-2} + A_3 \alpha^{-3}}{1 + B_1 \alpha^{-1} + B_2 \alpha^{-2}} \quad (\text{A4.59})$$

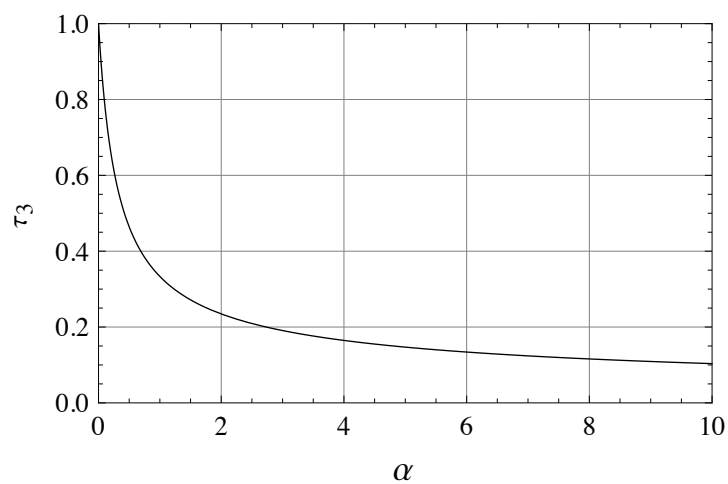
$$\tau_4 \approx \frac{C_0 + C_1 \alpha^{-1} + C_2 \alpha^{-2} + C_3 \alpha^{-3}}{1 + D_1 \alpha^{-1} + D_2 \alpha^{-2}} \quad (\text{A4.60})$$

$\alpha < 1$ の場合,

$$\tau_3 \approx \frac{1 + E_1 \alpha + E_2 \alpha^2 + E_3 \alpha^3}{1 + F_1 \alpha + F_2 \alpha^2 + F_3 \alpha^3} \quad (\text{A4.61})$$

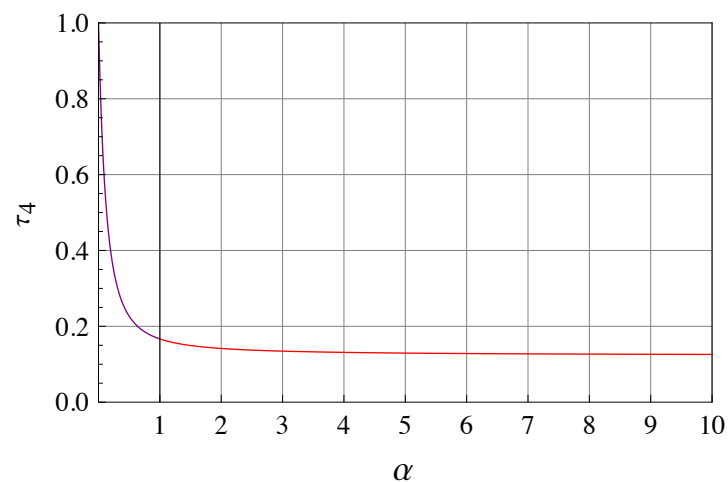
$$\tau_4 \approx \frac{1 + G_1 \alpha + G_2 \alpha^2 + G_3 \alpha^3}{1 + H_1 \alpha + H_2 \alpha^2 + H_3 \alpha^3} \quad (\text{A4.62})$$

式(A4.57)の α と τ_3 の関係を図A4.4に示す. 式(A4.59)と式(A4.61)はこの図とほとんど重なる.



図A4.4 Pearson type III 分布の α と τ_3 の関係 (式(A4.57)および式(A4.59とA4.61))

また, 式(A4.60)と式(A4.62)による α と τ_4 の関係を図A4.5に示す.



図A4.5 Pearson type III 分布の α と τ_4 の関係 (式(A4.60とA4.62))

式(A4.59)～(A4.62)の係数は表A4.2のとおりである.

母数推定法：

$$\alpha \approx \frac{1 + 0.2906 z}{z + 0.1882 z^2 + 0.0442 z^3}, \quad z = 3 \pi \tau_3^2, \quad \text{for } 0 < |\tau_3| < \frac{1}{3} \quad (\text{A4.63})$$

$$\alpha \approx \frac{0.36067 z - 0.59567 z^2 + 0.25361 z^3}{1 - 2.78861 z + 2.56096 z^2 - 0.77045 z^3}, \quad z = 1 - |\tau_3|, \quad \text{for } \frac{1}{3} \leq |\tau_3| < 1 \quad (\text{A4.64})$$

α が与えられる場合には次式の方がよい.

$$\gamma = \frac{2}{\sqrt{\alpha}} \text{sign}(\tau_3), \quad \sigma = \lambda_2 \sqrt{\alpha \pi} \Gamma(\alpha) / \Gamma\left(\alpha + \frac{1}{2}\right), \quad \mu = \lambda_1 \quad (\text{A4.65})$$

表A4.2 式(A4.59)～(A4.62)の係数

$A_0 = 3.2573501 \times 10^{-1}$	$C_0 = 1.2260172 \times 10^{-1}$
$A_1 = 1.6869150 \times 10^{-1}$	$C_1 = 5.3730130 \times 10^{-2}$
$A_2 = 7.8327243 \times 10^{-2}$	$C_2 = 4.3384378 \times 10^{-2}$
$A_3 = -2.9120539 \times 10^{-3}$	$C_3 = 1.1101277 \times 10^{-2}$
$B_1 = 4.6697102 \times 10^{-1}$	$D_1 = 1.8324466 \times 10^{-1}$
$B_2 = 2.4255406 \times 10^{-1}$	$D_2 = 2.0166036 \times 10^{-1}$
$E_1 = 2.3807576$	$G_1 = 2.1235833$
$E_2 = 1.5931792$	$G_2 = 4.1670213$
$E_3 = 1.1618371 \times 10^{-1}$	$G_3 = 3.1925299$
$F_1 = 5.1533299$	$H_1 = 9.0551443$
$F_2 = 7.1425260$	$H_2 = 2.6649995 \times 10^1$
$F_3 = 1.9745056$	$H_3 = 2.6193668 \times 10^1$

A-5 一級河川における流域平均雨量のAMS資料の特性

第4章4.1で用いた一級水系の観測地点，上流流域面積，観測年数，3時間，6時間，12時間および24時間の流域平均雨量のAMS資料にGEV分布をあてはめた際の形状母数を示す。

地域	水系名河川名	観測地点	上流流域面積 km ²	観測 年数	3-hour	6-hour	12-hour	24-hour	平均値
北海道	網走川水系網走川	美幌	825	45	-0.013	0.082	0.001	-0.062	0.002
	石狩川水系石狩川	石狩大橋	12697	44	-0.023	-0.027	-0.090	-0.199	-0.085
	釧路川水系釧路川	標茶	895	45	-0.062	-0.076	-0.007	0.033	-0.028
	鷲川水系鷲川	鷲川	1228	35	-0.002	-0.023	-0.112	-0.187	-0.081
	留萌川水系留萌川	大和田	234	30	-0.360	-0.357	-0.438	-0.451	-0.401
	沙流川水系沙流川	平取	1253	37	0.008	-0.022	-0.098	-0.077	-0.047
	後志利別川水系後志利別川	今金	361	38	-0.034	0.036	-0.042	-0.047	-0.022
	尻別川水系尻別川	名駒	1402	36	-0.036	-0.138	-0.225	-0.259	-0.165
	渚滑川水系渚滑川	上渚滑	1051	45	-0.028	0.068	0.089	-0.031	0.025
	天塩川水系天塩川	誉平	4029	30	-0.012	0.037	-0.046	-0.210	-0.058
	十勝川水系十勝川	茂岩	8277	36	-0.056	-0.006	-0.066	-0.199	-0.082
	常呂川水系常呂川	北見	1394	42	-0.178	-0.164	-0.139	-0.057	-0.135
	湧別川水系湧別川	開盛	1335	37	0.044	0.118	0.024	-0.115	0.018
		平均値		38.5	-0.058	-0.036	-0.088	-0.143	-0.081
東北地方	阿武隈川水系阿武隈川	岩沼	5265	41	-0.176	-0.213	-0.105	-0.098	-0.148
	岩木川水系岩木川	五所川原	1740	39	-0.122	-0.092	0.025	-0.072	-0.065
	北上川水系北上川	狐禅寺	7070	53	-0.103	-0.073	-0.051	-0.051	-0.069
	北上川水系旧北上川	和渚	1884	52	-0.261	-0.185	-0.103	-0.136	-0.171
	鳴瀬川水系吉田川	落合	197	50	-0.257	-0.270	-0.245	-0.256	-0.257
	鳴瀬川水系鳴瀬川	三本木	551	50	-0.224	-0.264	-0.234	-0.224	-0.237
	名取川水系名取川	名取橋	456	47	-0.137	-0.083	-0.107	-0.146	-0.118
	名取川水系広瀬川	広瀬橋	312	47	-0.111	-0.107	-0.152	-0.197	-0.142
	雄物川水系雄物川	椿川	4035	54	0.195	0.131	0.186	0.099	0.153
	高瀬川水系高瀬川	高瀬橋	850	41	-0.179	-0.122	-0.070	-0.050	-0.105
	米代川水系米代川	二ツ井	3750	48	-0.123	-0.027	0.167	0.147	0.041
		平均値		47.5	-0.136	-0.119	-0.063	-0.089	-0.102
関東地方	荒川水系荒川	治水橋	2015	67	-0.021	-0.011	-0.015	-0.068	-0.029
	利根川水系鬼怒川	石井	1230	61	0.166	0.217	0.194	0.101	0.170
	利根川水系小貝川	黒子	580	62	-0.189	-0.191	-0.185	-0.129	-0.174
	那珂川水系那珂川	野口	2181	44	-0.014	-0.047	-0.112	-0.147	-0.080
	相模川水系相模川	厚木	1211	39	0.105	0.163	0.128	0.041	0.110
	利根川水系利根川	八斗島	5150	60	-0.052	0.069	-0.052	-0.157	-0.048
	鶴見川水系鶴見川	末吉橋	230	50	0.214	0.063	-0.075	-0.072	0.032
	利根川水系渡良瀬川	高津戸	472	61	0.029	0.014	-0.026	-0.068	-0.013
		平均値		55.5	0.030	0.035	-0.018	-0.062	-0.004

北陸地方	荒川水系荒川	葛箆山	1079	40	-0.260	-0.299	-0.290	-0.249	-0.275
	阿賀野川水系阿賀野川	馬下	6997	39	-0.042	0.049	0.021	0.016	0.011
	阿賀野川水系阿賀川	山科	2742	38	0.003	0.005	-0.018	0.020	0.002
	阿賀野川水系只見川	片門	2765	38	-0.136	-0.018	-0.004	-0.030	-0.047
	信濃川水系信濃川下流	帝石橋	1260	41	-0.133	-0.075	-0.153	-0.185	-0.137
	関川水系関川	高田	703	41	0.096	0.061	-0.010	-0.196	-0.012
	関川水系保倉川	松本	332	41	-0.034	-0.044	-0.072	-0.051	-0.050
	姫川水系姫川	山本	698	38	0.016	-0.035	-0.170	-0.183	-0.093
	黒部川水系黒部川	愛本	667	41	0.105	0.144	-0.012	-0.015	0.055
	常願寺川水系常願寺川	瓶岩	344	44	-0.172	-0.130	-0.052	-0.027	-0.095
	神通川水系神通川	神通大橋	2688	39	-0.091	-0.027	0.131	0.111	0.031
	神通川水系井田川	高田橋	408	39	-0.052	-0.105	0.060	0.137	0.010
	神通川水系熊野川	八幡橋	128	39	-0.039	-0.014	0.032	0.039	0.004
	庄川水系庄川	雄神	1100	43	-0.028	0.053	0.052	0.079	0.039
	小矢部川水系小矢部川	津沢	280	45	-0.042	-0.058	-0.079	-0.054	-0.058
	手取川水系手取川	鶴来	748	47	0.060	0.042	0.065	0.178	0.086
	梯川水系梯川	埴田	155	31	0.095	-0.028	-0.044	0.070	0.023
		平均値		40.2	-0.038	-0.028	-0.032	-0.020	-0.030
中部地方	安倍川水系安倍川	手越	537	41	0.070	0.062	0.057	0.101	0.072
	狩野川水系狩野川	大仁	306	64	-0.041	-0.136	-0.146	-0.111	-0.108
	菊川水系菊川	国安	149	45	-0.058	0.080	0.020	-0.031	0.002
	鈴鹿川水系鈴鹿川	高岡	267	44	-0.069	-0.103	-0.086	0.004	-0.063
	雲出川水系雲出川	雲出橋	541	38	0.027	0.010	-0.019	0.061	0.020
	櫛田川水系櫛田川	両郡橋	379	38	0.156	0.101	0.152	0.207	0.154
	宮川水系宮川	岩出	780	39	0.222	0.200	0.125	0.107	0.163
	大井川水系大井川	神座	1,160	41	-0.028	0.007	0.193	0.300	0.118
	庄内川水系庄内川	枇杷島	731	37	0.112	-0.113	-0.130	-0.143	-0.069
	天竜川水系天竜川	宮ヶ瀬	2224	41	-0.178	-0.084	-0.055	-0.114	-0.108
	天竜川水系天竜川	鹿島	4,955	41	0.107	0.016	0.041	0.172	0.084
	豊川水系豊川	石田	545	42	-0.076	-0.038	0.054	0.217	0.039
	矢作川水系矢作川	岩津	1,375	41	-0.010	0.004	0.060	0.189	0.061
	木曽川水系木曽川	犬山	4,827	42	0.224	0.189	0.297	0.164	0.219
	木曽川水系長良川	忠節	1,622	43	-0.114	-0.012	0.089	0.112	0.019
	木曽川水系揖斐川	万石	1,196	42	0.125	0.116	0.163	0.015	0.105
		平均値		42.4	0.029	0.019	0.051	0.078	0.044
中国地方	旭川水系旭川	下牧	1588	43	-0.184	-0.067	0.014	0.101	-0.034
	芦田川水系芦田川	山手	799	39	-0.141	-0.053	0.067	0.132	0.001
	江の川水系江の川	川平	3836	48	-0.009	-0.041	-0.004	-0.026	-0.020
	江の川水系江の川	尾関山	2006	48	0.094	0.115	0.076	0.066	0.088
	斐伊川水系斐伊川	上島	895	42	0.144	0.136	0.174	0.110	0.141
	日野川水系日野川	車尾	857	42	0.240	0.240	0.104	0.107	0.173

	佐波川水系佐波川	新橋	423	42	-0.048	-0.016	0.182	0.229	0.087
	千代川水系千代川	行徳	1057	42	-0.086	-0.096	-0.029	0.009	-0.051
	高梁川水系高梁川	船穂	2639	43	0.011	0.153	0.177	0.136	0.119
	高津川水系高津川	高角	1080	44	0.270	0.360	0.156	-0.000	0.196
	天神川水系天神川	小田	464	42	0.067	-0.016	-0.104	-0.114	-0.042
	吉井川水系吉井川	河田原	1929	44	-0.121	-0.064	0.003	-0.041	-0.056
	小瀬川水系小瀬川	両国橋	330	32	-0.159	-0.167	0.047	0.359	0.020
	太田川水系太田川	矢口第一	1530	43	0.176	0.217	0.351	0.104	0.212
		平均値		42.4	0.018	0.050	0.087	0.084	0.060
四国地方	土器川水系土器川	祓川橋	107	39	-0.238	-0.195	-0.139	-0.084	-0.164
	那賀川水系桑野川	大原	69	43	-0.071	0.042	0.108	0.019	0.024
	那賀川水系那賀川	古庄	765	47	0.041	0.037	0.120	0.181	0.095
	物部川水系物部川	深淵	468	41	0.286	0.196	0.084	0.148	0.178
	重信川水系重信川	出合	445	53	0.007	-0.013	0.017	-0.050	-0.010
	重信川水系石手川	湯渡	105	53	-0.029	-0.025	-0.035	-0.010	-0.025
	吉野川水系吉野川	岩津	2810	44	0.031	0.073	0.155	0.256	0.129
	肱川水系肱川	大洲	1009	42	0.136	0.321	0.370	0.327	0.289
	仁淀川水系仁淀川	伊野	1463	43	0.064	0.081	0.129	0.066	0.085
	渡川水系中筋川	磯ノ川	94	43	0.300	0.079	0.002	-0.052	0.082
	渡川水系四万十川	具同	1808	43	0.247	0.212	0.178	0.022	0.165
	渡川水系後川	秋田	133	36	0.373	0.356	0.111	-0.017	0.206
		平均値		43.9	0.096	0.097	0.092	0.067	0.088
九州地方	番匠川水系番匠川	番匠橋	278	39	0.245	0.224	0.104	0.124	0.174
	五ヶ瀬川水系五ヶ瀬川	三輪	1044	44	0.127	0.066	-0.023	-0.061	0.027
	本明川水系本明川	裏山	36	41	-0.278	-0.292	-0.322	-0.247	-0.285
	嘉瀬川水系嘉瀬川	官人橋	226	45	0.063	-0.033	-0.053	-0.039	-0.016
	菊池川水系菊池川	玉名	906	45	-0.139	-0.087	0.095	0.189	0.015
	肝属川水系肝属川	俣瀬	450	60	0.226	0.212	0.128	0.208	0.194
	球磨川水系球磨川	横石	1856	45	0.062	-0.003	0.105	0.023	0.047
	松浦川水系松浦川	松浦橋	434	45	0.168	0.158	0.074	0.071	0.118
	緑川水系緑川	城南	681	64	0.065	0.022	0.091	0.049	0.057
	大淀川水系大淀川	柏田	2126	44	0.046	0.154	0.122	0.146	0.117
	川内川水系川内川	川内	1409	45	0.094	0.129	0.129	0.046	0.099
	筑後川水系筑後川	荒瀬	1443	55	-0.018	0.093	0.053	0.106	0.058
		平均値		47.7	0.055	0.054	0.042	0.051	0.050
		全国平均値		43.9	-0.001	0.008	0.011	0.002	0.005

参考文献

- Greenwood, J.A., J.M. Landwehr, N.C. Matalas and J.R. Wallis: Probability Weighted Moments —Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form, *Water Resources Research*, Vol.15, No.5, pp.1049-1054, 1979.
- Greis, N.P. and E.F. Wood: Regional flood frequency estimation and network design, *Water Resources Research*, Vol.17, pp.1167-77, 1981. (Correction: *Water Resources Research*, Vol.19, pp.589-90, 1983.)
- Hosking, J.R.M., and J.R. Wallis: Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, Vol.29, pp.339-49, 1987.
- Hosking, J.R.M., and J.R. Wallis: Regional Frequency Analysis –An Approach Based on L-Moments–, Cambridge Univ. Press., p.207, 1997.
- Hosking, J.R.M., J.R. Wallis, and E.F. Wood: An appraisal of the regional flood frequency procedure in the UK Flood Studies Report. *Hydrological Sciences Journal*, Vol.30, pp. 85-109, 1985a.
- Hosking, J.R.M., J.R. Wallis, and E.F. Wood: Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics*, Vol.27, pp.251-61, 1985b.
- Landwehr, J.M., N.C. Matalas, and J.R. Wallis: Probability-weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, Vol.15, pp.1055-64, 1979a.
- Landwehr, J.M., N.C. Matalas, and J.R. Wallis: Estimation of parameters and quantiles of Wakeby distributions. *Water Resources Research*, Vol.15, pp.1361-79, 1979b. (Correction: *Water Resources Research*, Vol.15, p.1672, 1979.)
- Stedinger, J.R.: Estimating a regional flood frequency distribution. *Water Resources Research*, Vol.19, pp.503-10, 1983.
- Wallis, J.R.: Risk and uncertainties in the evaluation of flood events for the design of hydraulic structures, In *Piène e Siccità*, edited by E. Guggino, G. Rossi, and E. Todini, pp.3-36. Fondazione Politecnica del Mediterraneo. Catania, Italy, 1981.
- Wallis, J.R.: Hydrologic problems associated with oilshale development, In *Environmental Systems and Management*, edited by S. Rinaldi, pp.85-102, 1982.
- Wallis, J.R., N.C. Matalas, and J.R. Slack: Just a moment!, *Water Resources Research*, Vol.10, 211-9, 1974.
- Wilkins, J. E.: A note on skewness and kurtosis. *Annals of Mathematical Statistics*, Vol.15, pp. 333-5, 1944.